DOCUMENT RESUME

ED 416 234 TM 028 114

AUTHOR	Brick, J. Michael; Collins, Mary A.; Nolin, Mary Jo; Davies, Elizabeth; Feibus, Mary L.
TITLE	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93). Working Paper Series.
INSTITUTION	Westat, Inc., Rockville, MD.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-WP-97-04
PUB DATE	1997-02-00
NOTE	114p.
AVAILABLE FROM	U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Avenue, N.W., Room 400, Washington, DC 20208-5654.
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Evaluative (142)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Data Collection; *Databases; Discipline; Early Childhood Education; Elementary School Students; *Interviews; Parents; Preschool Children; *Research Design; Research Methodology; School Readiness; School Safety; *Surveys; Tables (Data)
IDENTIFIERS	*Monitoring; *National Household Education Survey

ABSTRACT

The National Household Education Survey (NHES) is a data collection system of the National Center for Education Statistics that collects and publishes data on the condition of education in the United States. It is a telephone survey of the noninstitutionalized population of the country, and it focuses on issues that are best studied through contacting households rather than institutions. The primary purpose of the NHES is to conduct repeated measures of the same phenomena at different points in time. Full-scale implementations of the NHES have been conducted in 1991, 1993, 1995, and 1996. This paper presents information on the NHES:93 survey design, data collection procedures, monitoring of interviews, interview administration time, and data editing. The 1993 NHES addressed readiness for school and safety and discipline in school. The School Readiness component covered a number of experiences of preschool children that have a bearing on readiness to start school. This survey used interviews with parents to sample 10,888 children aged 3 through 7, or in second grade and below. The School Safety and Discipline component focused on school environment, school safety, school discipline policy, and alcohol and other drug use and education. For this component, parents of 12,680 children in grades 3 through 12. The following sections are included: (1) "Overview of the National Household Education Survey"; (2) "Monitoring Interviews in the NHES:93"; (3) "Survey Administration Time for the NHES:93"; and (4) "NHES:93 Data Editing." Three appendixes contain a monitoring form, 22 graphs of monitoring results, and a database design diagram. (Contains 3 figures, 12 tables, and 17 references.) (SLD)

ED 416 234

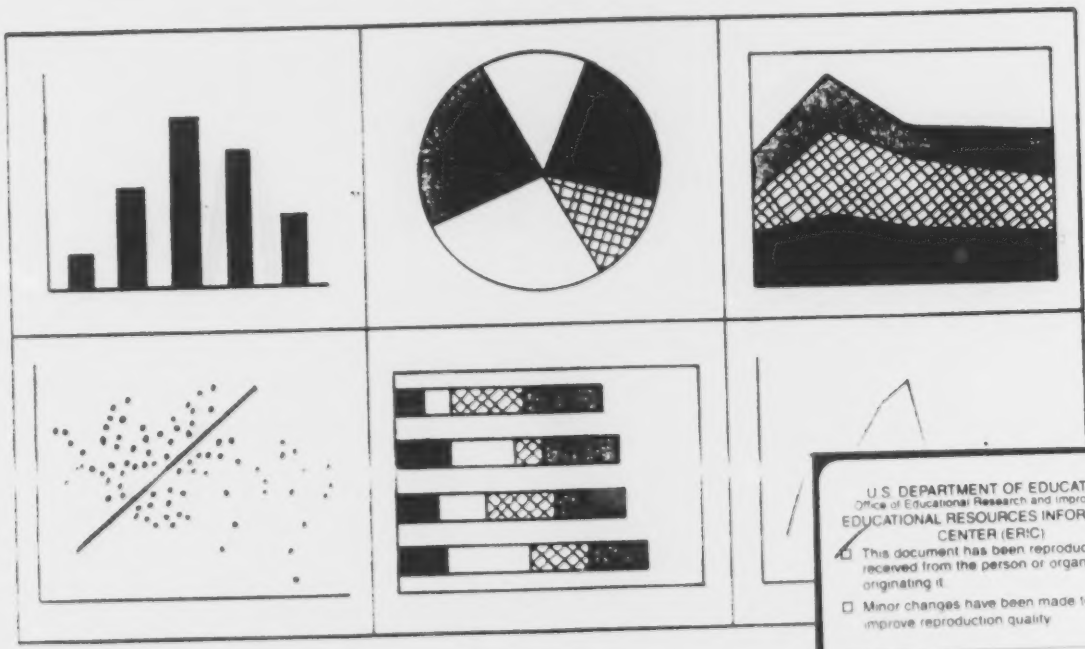
NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)

Working Paper No. 97-04

February 1997



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

U.S. Department of Education
Office of Educational Research and Improvement

BEST COPY AVAILABLE

Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)

Working Paper No. 97-04 – February 1997

Contact: Kathryn Chandler
Surveys and Cooperative Systems Group
(202) 219-1767
e-mail: nhes@ed.gov
www.ed.gov/NCES/NHES

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Marshall S. Smith
Acting Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Surveys and Cooperative Systems Group
Paul D. Planchon
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)*, Working Paper No. 97-04, by J. Michael Brick, Mary A. Collins, Mary Jo Nolin, Elizabeth Davies, and Mary L. Feibus. Project Officer, Kathryn Chandler. Washington, D.C.: 1997.

February 1997

Foreword

Each year a large number of written documents are generated by NCES staff and individuals commissioned by NCES which provide preliminary analyses of survey results and address technical, methodological, and evaluation issues. Even though they are not formally published, these documents reflect a tremendous amount of unique expertise, knowledge, and experience.

The ***Working Paper Series*** was created in order to preserve the information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series. Consequently, we encourage users of the series to consult the individual authors for citations.

To receive information about submitting manuscripts or obtaining copies of the series, please contact Ruth R. Harris at (202) 219-1831 or U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave., N.W., Room 400, Washington, D.C. 20208-5654.

Susan Ahmed
Chief Mathematical Statistician
Statistical Standards and
Services Group

Samuel S. Peng
Director
Methodology, Training, and Customer
Service Program

Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)

Prepared by:

J. Michael Brick
Mary A. Collins
Mary Jo Nolin
Elizabeth Davies
Mary L. Feibus

Westat, Inc.

Prepared for:

U.S. Department of Education
Office of Educational Research and Development
National Center for Education Statistics

February 1997

Table of Contents

Page

Foreword ... iii

Overview of the National Household Education Survey ... 1

NHES:93 Design ... 2

Data Collection Experience ... 13

Supervisor and Interviewer Staff Training ... 13
Data Collection Procedures ... 14
Data Quality Control Activities ... 22

Monitoring Interviews in the NHES:93 ... 24

Monitoring Form ... 24
Monitoring Sessions ... 24
Data Reliability ... 25
Patterns Over Time ... 26
General Patterns ... 27
Interview Type ... 28
Summary ... 36

Survey Administration Time for the NHES:93 ... 38

Editing the Administration Time ... 38
Screener Administration Time ... 39
School Readiness Administration Time ... 40
School Safety and Discipline Administration Time ... 41

NHES:93 Data Editing ... 45

Data Alignment ... 45
Range Edits ... 45
Logic Edits ... 46
Structural Edits ... 46
Frequency and Cross-Tabulation Review ... 46
Frequency Review of Text Items ... 46
Problem Areas and Suggestions for Improvements in Future Surveys ... 47
Logic and Range Edit Specifications ... 48
Edits for Structural Completeness ... 58
Skip Edits ... 63

References ... 65

Table of Contents–Continued

Page

Appendices

Appendix A: Monitoring Form A-1
Appendix B: Graphs of Monitoring Results B-1
Appendix C: Database Design Diagram C-1

Tables

1. Percentage of interviewer hours and completed interviews, per week, NHES:93 16
2. Results of Screener refusal conversion efforts on the NHES:93 19
3. Results of extended interview refusal conversion efforts on the NHES:93 19
4. Results of refielding NHES Screeners in maximum calls, NHES;93 20
5. Number of monitoring sessions, by length of session 25
6. Number of monitoring sessions, by week 26
7. Summary of reading and general skills from monitoring forms, by type of interview 31
8. Summary of listening and probing skills from monitoring forms, by type of interview . . . 32
9. Summary of recording skills from monitoring forms, by type of interview 33
10. Summary of handling refusals skills from monitoring forms, by type of interview 34
11. Summary of telephone manner skills from monitoring forms, by type of interview 35
12. Mean, median, and quartile administration time (in minutes) of NHES:93 completed interviews, by interview type and segment 43

Figures

1. Cumulative number of interviewer hours and completed interviews, by week 17
2. Cumulative number of completed interviews by type of interview, by week 17
3. Mean administration time in minutes of Screener over data collection period 40

Overview of the National Household Education Survey

The National Household Education Survey (NHES) is a data collection system of the National Center for Education Statistics (NCES), which has as its legislative mission the collection and publication of data on the condition of education in the Nation. The NHES is specifically designed to support this mission by providing information on those educational issues that are best addressed by contacting households rather than schools or other educational institutions. The NHES provides descriptive data on the educational activities of the U.S. population and offers policymakers, researchers, and educators a variety of statistics on the condition of education in the United States.

The NHES is a telephone survey of the noninstitutionalized civilian population of the U.S. Households are selected for the survey using random-digit-dialing (RDD) methods, and data are collected using computer-assisted telephone interviewing (CATI) procedures. 45,000 to 60,000 households are screened for each administration, and individuals within households who meet predetermined criteria are sampled for more detailed or extended interviews. The data are weighted to permit estimates of the entire population. The NHES survey for a given year typically consists of a Screener, which collects household composition and demographic data, and extended interviews on two substantive components addressing education-related topics. In order to assess data item reliability and inform future NHES surveys, each administration also includes a subsample of respondents for a reinterview.

The primary purpose of the NHES is to conduct repeated measurements of the same phenomena at different points in time. Throughout its history, the NHES has collected data in ways that permit estimates to be tracked across time. This includes repeating topical components on a rotating basis in order to provide comparative data across survey years. In addition, each administration of the NHES has benefited from experiences with previous cycles, resulting in improvements to the survey procedures and content. Thus, while the survey affords the opportunity for tracking phenomena across time, it is also, dynamic in addressing new issues and including conceptual and methodological refinements.

A new design feature of the NHES program implemented in the NHES:96 was the collection of demographic and educational information on members of all screened households, rather than just those households potentially eligible for a topical component. In addition, this expanded screening feature was designed to include a brief set of questions on an issue of interest to education program administrators or policymakers. The total Screener sample size is sufficient to produce state estimates of household characteristics for the NHES:96.

Full-scale implementations of the NHES have been conducted in 1991, 1993, 1995, and 1996. Topics addressed by the NHES:91 were early childhood education and adult education. The NHES:93 collected information about school readiness and school safety and discipline. The 1991 components were repeated for the NHES:95, addressing early childhood program participation and adult education. Both components underwent substantial redesign to incorporate new issues, develop new measurement approaches, and reflect methodological advancements. In the NHES:96, the topical components were parent/family involvement in education and civic involvement. The NHES:96 expanded screening feature included a brief set of questions on public library use.

In addition to its topical components, the NHES system has also included a number of methodological investigations. These have resulted in technical reports and working papers covering diverse topics such as telephone undercoverage bias, proxy reporting, and sampling methods. This series of technical reports and working papers provides valuable information on ways of improving the NHES.

This working paper presents information on the NHES:93 survey design, data collection procedures, monitoring of interviews, interview administration time, and data editing. Readers may also wish to review *Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1993 National Household Education Survey* (Brick et al. forthcoming), and *Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey* (Brick et al. forthcoming) for additional information on the survey. Comparable working papers are being prepared for the NHES:95 and the NHES:96.

NHES:93 Design

The 1993 National Household Education Survey (NHES:93) addressed readiness for school and safety and discipline in school. These topics are related to Goal 1 and Goal 6, two of the National Education Goals. Specifically, Goal 1 states that "By the year 2000, all children in America will start school ready to learn." Goal 6 states that "By the year 2000, every school in America will be free of drugs and violence and will offer a safe, disciplined environment conducive to learning."

The School Readiness (SR) component covered experience in early childhood programs, the child's accomplishments and difficulties in several developmental domains, school adjustment and related problems, delayed kindergarten entry, and early primary school experiences, including repeating grades, the child's general health and nutritional status, home activities, and family characteristics such as stability and economic risk factors. Altogether, 10,888 children aged 3 through 7 on December 31, 1992, or in 2nd grade or below (up to a maximum age of 9) were sampled. Interviews were conducted with 4,423 parents of preschool children, 2,126 parents of kindergartners, 4,277 parents of primary school children, and 62 parents of home school children. For further information on the content of the SR component, see *National Household Education Survey of 1993: School Readiness Data File User's Manual* (Brick et al. 1994a).

The School Safety and Discipline component (SS&D) focused on four areas: school environment, school safety, school discipline policy, and alcohol/other drug use and education. The SS&D interview gathered general perceptions of the school learning environment from both parents and students. Parents of 12,680 children in 3rd through 12th grades (up to age 20 on December 31, 1992) were interviewed, as were 6,504 students in 6th through 12th grades. For further information on the content of the SS&D component, see *National Household Education Survey of 1993: School Safety and Discipline Data File User's Manual* (Brick et al. 1994b).

The NHES:93 was developed to provide reliable estimates for each of the two different components described above. The inclusion of two survey components made the overall survey more cost effective, thus allowing for larger sample sizes and more precise estimates. This strategy was key to the NHES design. By including more than one topic within the framework of a single survey, the cost of screening households to find those eligible for the study could be partitioned over the component surveys.

It was possible that the same household member could be selected to respond to more than one interview and/or that more than one household member could be sampled. For the SR interview, if there were one or two eligible children in the household, interviews were conducted for those children. If the household included more than two eligible children, two children were randomly sampled from that household. For the SS&D interview, if a household had one eligible youth, that youth was selected with a probability that depended on his/her grade (students in grades 3 through 5 were selected with a lower probability than those in grades 6 through 12). If a household had two or more eligible youths, the sampling depended upon the number of youths in the household in each of the two grade categories. A maximum of

two youths was selected from any household for the SS&D component, one from the lower grades and one from the upper grades.

Even though sampling methods reduced the number of interviews per household, the length of the interview was considered to be a critical factor in obtaining high response rates and reliable estimates. Therefore, the number of items included in the NHES:93 was limited in order to help improve response rates and reduce the demands made on survey respondents.

Because of the above requirements, complex sampling techniques, and the need for quick and accurate administration, the NHES:93 was conducted using computer assisted telephone interviewing (CATI) technology. Some of the advantages of CATI for the NHES:93 included improved project administration, online sampling and eligibility checks, scheduling of interviews according to a priority scheme to improve response rates, managing data quality by controlling skip patterns and checking responses online for range and consistency, and an online "help" function to answer interviewers' questions.

Three different interview instruments were used in the NHES:93. These instruments were the Screener, the SR interview, and the SS&D interview, which included specific sets of questions for parents and for youth. Items within each of the instruments were programmed so that the appropriate items appeared on the interviewer's computer screen corresponding to the respondent's answer to previous queries.

Random-Digit-Dialing Sampling

The sampling method used for the NHES:93 was a variant of RDD procedures described in Waksberg (1978). The original Mitofsky-Waksberg method produces an equal probability sample of households with telephones and requires a smaller number of telephone calls than the sampling procedures previously used for RDD. A time-saving variant of this method, referred to as the modified Waksberg procedure, was used for the NHES. The modified method is described in Brick and Waksberg (1991).

The basic operating method of RDD sampling is simple. A list of all existing telephone area codes and prefix numbers is determined for the 50 States and the District of Columbia. (The prefix numbers are the three-digit telephone exchanges.) All possible combinations for the next two digits are added to the set of prefix numbers. Thus, a list is established of all the possible first 8 digits of the 10 digits in telephone numbers. These eight digit numbers are treated as Primary Sampling Units (PSUs), or telephone clusters.

A random sample of eight-digit PSUs is selected. A "prime" telephone number is formed by adding a random two-digit number to the eight-digit cluster. The prime number is then dialed. If the prime number is residential, the PSU is retained in the sample. If the prime number is not residential, then the PSU is rejected and no further calls within the PSU are made. Additional PSUs are selected in the same way until a predesignated number of PSUs is chosen.

A random sample of telephone numbers within each of the retained "residential" PSUs is selected by adding random two-digit combinations to the original eight numbers. Interviews are attempted at the prime number and at as many additional numbers required to obtain the desired within PSU sample size. The total expected sample size is $m(k+1)$, where m is the number of residential PSUs and $k+1$ is the number of households needed in each PSU. The values of m and k are chosen to satisfy the criterion for an optimum sample design. This sampling procedure produces a sample of telephone households with equal probabilities.

In the Mitofsky-Waksberg procedure, the households are sampled within clusters rather than in a simple random sample of households. The cost of the cluster sampling is much lower than that of simple random sampling of households because there are considerable savings in the telephone operations. With cluster sampling, the reduction in the number of telephone numbers to be dialed is at least 50 percent. However, the variances of the estimates are increased slightly due to the clustering of the sampled households with the PSUs. This variance increase is discussed below.

The original Mitofsky-Waksberg method is a sequential process. It is not possible to determine in advance how many telephone numbers in a cluster need to be dialed in order to achieve a sample size of k+1 households, and a new determination needs to be made for a PSU after each telephone call. This is awkward to implement when a survey faces a tight deadline, as was the case with the NHES:93.

An alternative, and much faster, sampling method is to use a fixed number of telephone numbers per PSU, rather than a fixed number of households. This method also provides a probability sample, but it is no longer self-weighting. This is the modified Waksberg procedure that was used for NHES:91 and again in 1993. A sample that is not self-weighting does increase the sampling errors of the estimates. Brick and Waksberg (1991) showed that the variance increases due to the departure from a self-weighting sample using the modified method are expected to be less than 10 percent using the procedures for the NHES:93.

Sampling Clusters

As mentioned above, the RDD sampling procedure is a two-stage sample design. The first stage involves the selection of PSUs or clusters of 100 numbers, and the second stage is a sample of telephone numbers within clusters. For a given sample size, it is possible to use any number of clusters (within certain limits), but the number of clusters affects both the cost of data collection and the precision of the estimates. The more clusters used, the higher the cost of data collection. However, using more clusters increases the precision of the estimates. A compromise between the cost and precision requirements is needed. The compromise involves taking enough clusters so that the precision is not greatly reduced and, at the same time, most of the savings of cluster sampling apply. The issues associated with the choice of the number of clusters and interviews within cluster are examined below.

The variance of an estimated mean from a cluster sample, selected in the way described above, can be approximated as:

$$\sigma^2(\bar{x}) = \frac{\sigma^2}{mb}(1+\rho(b-1))$$

where

σ^2 = the population variance for the characteristic, $\bar{x}$,

m = number of clusters in the sample,

b = average number of interviews completed per cluster, and

ρ = intraclass correlation for x.

The intraclass correlation, ρ, is a measure of the extent to which persons within small geographic areas (in this case, represented by 100 consecutive telephone numbers) are more similar to each other in regard to the estimated characteristics than to the general population. There is a limited body of research on the size of intraclass correlation coefficients in RDD surveys, but most of this research indicates that the correlations tend to be small. Groves and Kahn (1979) estimated the average intraclass correlation coefficient for a set of attitudinal items in their study at about 0.08. Alexander et al. (1986) estimated intraclass correlations for some health-related characteristics and found that they ranged from about 0.001 to 0.03. For education and age, they estimated intraclass correlations to be larger, ranging from 0.10 to 0.15.

Alexander et al. (1986) also estimated the optimal number of completed interviews per cluster (b) for statistics with different average intraclass correlations, based on the cost structure observed in their survey. The optimal number of interviews per cluster ranged from 40 (for ρ=0.001) to 3 (for ρ=0.15). In a discussion published with Alexander et al. (1986), Waksberg noted that their cost structure was very different from that observed in Westat studies. The cost structure found in Westat studies indicated that, for equal intraclass correlations (ρ), smaller values of b than proposed by Alexander et al. were optimal.

In subsequent sections, the average number of interviews per cluster are presented for the School Readiness and SS&D components of the NHES:93. The value of b is the number of completed interviews for the characteristic of interest, not the number of households. If only a subset of the completed interviews are used to estimate the characteristic (only children enrolled in first grade, for example), then the value of b is based on the number of respondents in that subset. As we shall point out, the value of b was less than 3 for most statistics produced from the NHES:93.

Sampling Households

For the NHES:93, screening interviews were completed with 64,708 households. The number of households needed for the survey (a target of 64,000) was based upon precision requirements for the estimates, in particular the requirements of the School Readiness component. The reasons for choosing this size sample are discussed in the next section.

The households were selected from 4,000 residential clusters, in other words, using the previous terminology, m equals 4,000 clusters. The average number of households screened per cluster was 16 (64,000÷4,000). This number determined the number of telephone numbers that had to be sampled within each cluster. Since many of the screened households did not have any household members who were of the appropriate ages, not all screened households were eligible for the survey. Thus, the average number of completed interviews per cluster was much smaller than 16.

The number of telephone numbers that had to be dialed in order to achieve the target sample of 64,000 households was estimated in steps. Since it takes approximately five phone numbers to identify a household cluster, it required about 20,000 phone numbers to identify a sample of 4,000 residential clusters (5 times 4,000). Each cluster identified as residential was kept regardless of whether a screener was completed for that household. Approximately 80 percent of these 4,000 residential contacts resulted in a completed screener. The second step involved locating residential numbers within the sampled clusters. In general, about 62 percent of random numbers dialed within a residential cluster under a modified Waksberg design are households. Assuming a response rate of 80 percent to the household screening questions, then the number of additional telephone numbers needed within the sampled clusters would be about 123,000 (60,800 additional completed households within the clusters divided by the .62 residency rate equals 98,065 numbers and 98,065 numbers divided by the .80 response rate equals 122,580 numbers). To ensure achieving this sample size, 31 additional telephone numbers were included for each cluster (124,000

telephone numbers instead of 122,580). Thus, the overall number of telephone numbers per cluster was 32, and the total number of phone numbers needed for both steps was 144,000 (124,000 plus 20,000). The actual completion of 64,708 screeners reflects a screening response rate slightly higher than the projected 80 percent (survey response is addressed in another working paper, *Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1993 National Household Education Survey* (Brick et al. forthcoming).

Within Household Sampling

In the previous sections, the procedures that were used to sample clusters and households for the NHES:93 were described. The following sections discuss sampling of persons within those households for the School Readiness and SS&D interviews and the reasons for choosing 64,000 households as the target sample size for the NHES:93. The School Readiness within-household sampling is described first.

School Readiness Interviews

The interviews for the School Readiness component were conducted with the parents of children 3 years old or older and who are not yet in the third grade and any children seven or younger regardless of grade. If there were one or two eligible children in a household, an interview was conducted for each of those children. If there were more than two eligible children in the household, two children were randomly sampled from that household. This sampling procedure served to limit the burden for the parents who have more than two children and did not deviate significantly from an equal probability sampling procedure.[1]

Estimates from the March 1991 Current Population Survey (CPS) indicated that the percent of all households with exactly one 3- to 7-year-old child was 11.8 percent, and the percent of all households with two or more 3- to 7-year-olds was 4.5 percent.[2] Therefore, the expected number of children sampled for an interview in the 64,000 screened households was about 13,300 (64,000 times {.118 + 2 x .045} equals 13,312). Assuming a School Readiness interview completion rate of about 94 percent[3], the expected number of completed interviews for the School Readiness component was about 12,500. The actual number of completed interviews was 10,888 because the SR response rate was 90 percent, rather than the expected 94 percent.

With 10,888 completed interviews in 4,000 clusters, the value of b is 2.7. This is the largest value b can become, because it assumes that all the sampled children are included in the same analysis. More commonly, groups of children are analyzed separately. For example, estimates by specific year of age or by grade in school have values of b less than unity. The size of these values of b indicate that a larger sample of clusters would not have been optimal for the School Readiness component, since 2.7 is less than the smallest optimal value of b suggested by Alexander et al (1986).

The number of screened households in the NHES:93 was determined by the School Readiness component precision requirements. The key precision requirement for the School Readiness component was the ability to detect differences between estimated percentages of children in different subdomains as

[1]The impact of differential sampling rates on the variance of the estimates in this paper are computed using the procedures discussed in Waksberg (1973). In this case, the increase in the variance of the estimates due to unequal sampling rates is less than one percent. In other words, the design effect due to this factor is less than 1.01. Kish (1965, page 429) gives a good description of the general problem associated with using unequal sampling rates in surveys.

[2]The estimated percent of households with 3- to 7-year-olds from the March CPS supplement was increased to account for the older children (those 8 years old or older) who will still be in second grade at the time of the survey.

[3]The completion rate for the Early Childhood Education component in the NHES:91 was 94.5 percent.

defined by the race and ethnicity of the child. For example, the comparison of the estimated percentage of children who were in a Head Start program prior to kindergarten by race and ethnicity is an important statistic. If the estimated percentage of black children with this characteristic differs by more than 15 or 20 percent from the estimated percentage of Hispanic children, then the sample size of the survey would be large enough to detect this difference at the 95 percent confidence level.

In the School Readiness survey, estimates for children entering school (kindergarten and first grade students) are especially important. With 64,000 screened households yielding approximately 12,500 completed interviews for 3- to 7-year old children and older children in second grade or below, we anticipated about 2,500 completed interviews for kindergarten students and approximately the same number of interviews for first grade students. Because of the oversampling procedure described in a later section, this target included about 300 completed interviews for both blacks and Hispanics in kindergarten. This sample would be large enough to detect relative differences of 20 percent or greater at the 95 percent confidence level.[4] The somewhat smaller number of completed interviews than expected means that the estimates are slightly less precise for these small subgroups. Of course, analyses of larger subgroups could identify smaller relative differences.

School Safety and Discipline Parent Interviews

Estimates from the March 1991 CPS data indicated that about 23,600 youths enrolled in the 3rd to the 12th grade would be identified in a sample of 64,000 households. This number exceeded the number of completed interviews needed to achieve the precision requirements for this component. Therefore, sampling within the screened households was done to obtain desired sample sizes and reduce the burden on respondents.

In the SS&D component, the sample size goals varied by the grade of the child. We begin by discussing the sample size requirements for the interviews with the parents of the students. The sample sizes for the interviews with the youths are discussed later.

For the youngest children, the sample size had to be large enough to support analysis of aggregates for those enrolled in the 3rd, 4th, and 5th grades. The safety and discipline issues addressed for these students were a subset of those addressed for the older students. A sample of approximately 2,600 students in these three grades was determined to be sufficient to meet the precision requirements of having a relative standard error (the standard error divided by the estimate) of 5.5 percent for a 40 percent characteristic.

For older youth in grades 6 through 12, the sample size needed to be large enough to support analysis of two-grade aggregates rather than three-grade aggregates. The comparison of the youth responses to the parent responses, which would only be done for the older youths, also required a larger sample size for students in grades 6 through 12. To achieve the precision requirements for the older students, an average two-grade aggregate sample size of about 3,000 completed interviews was established. This sample size provided for a relative standard error of 5.5 percent for a 40 percent characteristic of 11th and 12th grade students. The reason for increasing the sample size from 2,600 to 3,000 is to account for the reduction in the precision due to within-household sampling procedures described below.

---

[4]Sample sizes of 300 or more in each group are large enough to detect relative differences of 20 percent or more for estimates at the 95 percent confidence level. This calculation assumes that at least 40 percent of the subdomain has the characteristic. In other words, if 50 percent of blacks were estimated to have a characteristic and 60 percent of Hispanics were estimated to have the same characteristic, then the estimated relative difference (equal to 20 percent since 100{0.5-0.6}/ 0.5=20 percent) should be statistically significant at the 95 percent confidence level.

The targeted number of completed interviews with parents of students in the 3rd, 4th, and 5th grades was 2,600. The planned number of completed interviews with parents of the older students, those in grades 6 through 12 was about 10,700 (7 grades times about 1,525 per grade). Therefore, the total desired number of completed parent interviews for the SS&D component was 13,300.

To obtain completed interviews with the parents of 13,300 3rd to 12th graders, sampling was done by the grade of the child and the number of children in the household. Only students currently enrolled[5] in regular school in the appropriate grades were eligible for the survey. Assuming a completion rate of about 92 percent[6] for the interviews with the parents of the youths, a sample of about 14,500 youths needed to be selected to obtain the desired goal. The actual number of completed parent interviews was 2,563 for 3rd through 5th graders and 10,117 for 6 through 12th graders, reflecting a slightly lower-than-expected parent response rate. Survey response is discussed in detail in a separate working paper, *Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1993 National Household Education Survey* (Brick et al. forthcoming).

The sample of youths from the 64,000 screened households could be selected in a variety of ways. One approach would be to sample each youth with equal probability, irrespective of the number of eligible youths in the household. While this method would be efficient from a statistical perspective (in that the variances of the estimates would not be inflated by unequal probabilities of selection), it would result in multiple youths being selected from the same household. In fact, this method would require selecting at least three or four youths in some households. This is a serious disadvantage because it places a large burden on households with many eligible youths and may result in higher nonresponse. In addition, this approach would not satisfy the subsampling requirements for youths in the lower grades (3rd to 5th) and youths in the higher grades (6th to 12th).

A better alternative was to restrict the sample so that no-more than two youths were selected from the same household. If a household has one eligible youth, that youth was be selected with a probability that depends on his/her grade in school (students in the 3rd, 4th, and 5th grades were selected with a lower probability than those in higher grades). If a household had two or more eligible youths, then the sampling depended upon the number of youths in the household in each of the various grade categories. A maximum of two youths were selected from any household. In households with youths in both the lower and the upper grades, a maximum of one youth from each grade level was selected.

The actual number of youths sampled for parent interviews depended upon the number of youths in the various grade categories in the household. If there was one youth in the lower grades in the household, the youth had a probability of selection equal to 0.45. In households with two youths in the lower grades, exactly one youth was sampled in 90 percent of the households and no youths in 10 percent of the households. In households with more than two youths in the lower grades, exactly one youth was sampled from each household, and each had the same chance of being sampled.

The same type of procedure was independently applied for the higher grades, except the probabilities of selection were different. A youth in a household with only one youth in the higher grades was sampled. For households with two or more youths in the higher grades, one or two youths was sampled

---

[5]We considered the possibility of including youth who had dropped out of school during the past year in the survey, but the sample size for this group would have been less than 200 and the data collection instrument would have had to have been modified significantly to be appropriate for these youth.

[6]The 92 percent completion rate is based upon the 97 percent completion rate observed in the 1989 Field Test of the NHES for the Household Respondent Interview for youth aged 14 to 21 years. In that interview, any knowledgeable household respondent could complete the interview. Since the parent interview for the SS&D survey has a more stringent requirement, the expected completion rate has been lowered to 92 percent.

with equal probability from each household. Two youths were sampled from the higher grades, if, and only if, there were no youths in the lower grades in that household.

With this sampling plan, youths in households with multiple youths in a grade category had a different probability of being sampled than youths in households with only one eligible student. The sampling weights were be adjusted for these unequal probabilities of selecting youths within households to make the estimates unbiased. However, the variances of the estimates were slightly increased over an equal probability design because of the variability in the number of eligible youths in households. The impact of the differential sampling rates for the lower grade categories was insignificant, i.e., less than 1 percent. The impact was small because, as in the School Readiness component, the percent of students subject to rates that are different from the overall base rate (0.45) is small. Weighting and estimation procedures are discussed further in another working paper, *Unit and Item Response Rates, Weighting and Imputation Procedures in the 1993 National Household Education Survey* (Brick et al. forthcoming).

This situation was different for the youths in the higher grade levels. Only about 47 percent of the students were in households with only one eligible youth and the remainder were in households with more than one eligible youth. These youths were sampled with lower probabilities (for example, with three youths in the household, each of the youths had a probability of selection of 0.33 if there were also lower grade youths in the household and a probability of .67 if there were no lower grade youths in the household). The overall impact from the differential rates applied in this age group was an increase in the variance of the estimates of about 16 percent. Because of the variance increase, the sample sizes for the older students were raised for two grade aggregates from 2,600 to 3,000 completed interviews.

The approach of sampling up to two youths per household for parent interviews was used in the SS&D component. This scheme limited the number of parent interviews to a maximum of four (two School Readiness interviews and two SS&D interviews), but for the vast majority of households (over 99 percent) this maximum was not achieved. The loss in precision of the estimates due to unequal probabilities of selection resulting from variation in the number of eligible youths per household was insignificant for the lower grades and was larger, but acceptable, for the higher grades. Furthermore, this method provided control on the distribution of the sample by the grade level of the youths.

The impact of clustering of the telephone numbers in the 4,000 clusters can also be seen to be small, given the sample sizes for the lower and higher grades. Even assuming that all completed interviews for the students in the higher grades are analyzed together, the average number of completed parent interviews per cluster (b) is only 2.7. It is even smaller for analyses conducted on subsets of the youths.

School Safety and Discipline Youth Interviews

The last within-household sampling issue for the SS&D component was the sampling of 6th through 12th graders for youth interviews. For the most part, these youth were not younger than 11 years of age.

The interviews with students were only conducted for those youth in grades 6 through 12 with completed parent SS&D interviews. This allowed the parent the opportunity to give an informed consent to the interviewing of their child (since the issues covered in the youth interview were included in the parent interview). This protocol was also consistent with the objective of comparing youth and parent responses.

The interviewing of youth under the age of 14 years in household surveys has not been attempted very often, although such interviews in a school setting are more common. One notable exception is the 1990 Survey of Children and Parents. This was a national telephone survey of 1,738

parents and included 929 interviews of children age 10 to 17 living in the same households.[7] Interviewing the children over the telephone did not pose significant problems in that survey; however, there was some difficulty contacting the older children at home in order to interview them.

The Youth Risk Behavior Survey (YRBS), conducted as a supplement to the 1992 National Health Interview Survey (NHIS) was another household survey in which children were interviewed. In this survey, interviews with persons 12 to 21 years old were conducted in the respondent's home. The YRBS covered many highly sensitive issues, including behaviors that result in intentional or unintentional injuries, drug and alcohol use, tobacco use, sexual behaviors, and sexually transmitted diseases. One of the major recommendations from the research on the YRBS was that the questionnaire be revised substantially to ease the cognitive burden on the young respondents and to decrease response errors. In the development of the SS&D instrument for the NHES:93, these same issues were analyzed in our cognitive laboratory work.

Other key recommendations for the YRBS involved privacy and the reading ability of the respondents. Reading ability was not a problem in the NHES:93 because the questions were asked on the telephone. However, the privacy of the young respondent was an issue, even on the telephone. One concern was that the parent may listen on a telephone extension, limiting the truthfulness of the youth in their responses.

The SS&D component presented fewer problems for the privacy of the young respondents because the questions were less sensitive than those in surveys such as the YRBS and, in general, did not ask about the youth's behavior. Since the parents had already completed the interview, they were aware of the nature of the questions and possibly less inclined to listen to their children's responses. Also, we attempted to construct the response categories so that the youth could answer with a neutral response (yes or no, agree or disagree) rather than a more specific response that might arouse concern from a parent listening to the youth.

We also obtained two measures of the privacy of the interview. One was the interviewer's opinion about whether or not another person was listening to the interview. A second measure was a question posed to the youth at the end of the interview that asked whether or not the youth felt that he or she was able to answer privately. While these items do not improve the quality of the responses, they provide a means for analysts to compare the response distributions depending on some measures of the privacy of the interview.

The target sample was 7,600 students enrolled in grades 6 through 12 and was expected to yield about 6,800 completed interviews, assuming a completion rate of 90 percent.[8] The youth were sampled with equal probability (0.71) from the completed interviews of parents of 6th through 12th graders. This sample size was designed to provide estimates of the required precision for aggregates of three grade levels. In other words, the target sample size for students in grades 6 through 8 was 2,900, and this sample would include over 300 blacks and 300 Hispanics in these grade groups. The precision requirement for 300 black and Hispanic students is described earlier in the discussion of the School Readiness component.

---

[7]See the report *Speaking of Kids*, available from the National Commission on Children, 1111 Eighteenth Street, NW Suite 810, Washington DC 20036.

[8]The completion rate for the 1989 Youth Interview in the NHES Field Test was about 86 percent, including out-of-household youth selected via multiplicity sampling. The youth in the 1993 SS&D component were not sampled using multiplicity sampling, so that the loss associated with following out-of-household youth was not a consideration in the expected 1993 youth completion rate.

Oversampling Minorities

One of the goals of the NHES:93 is to produce reliable estimates for subdomains defined by race and ethnicity. In fact, estimates by race and ethnicity were key in developing the sample sizes for each of the components of the NHES:93. In a 64,000 household design in which every household has the same probability of being included, the number of completed interviews would not be large enough to produce reliable estimates of many characteristics of blacks and Hispanics. Therefore, blacks and Hispanics must be sampled at higher rates to improve the reliability of estimates for these domains.

In NHES:91 and NHES Field Test of 1989, we examined and used a particular method of oversampling blacks and Hispanics. The method employed was successful in reducing the variances for estimates of characteristics of blacks and Hispanics by approximately 20 to 30 percent over the range of statistics examined. The decreases in precision for estimates of the groups that were not oversampled and for estimates of totals were modest, ranging from about 5 to 15 percent. The results of using this method of oversampling are described by Mohadjer and West (1992). Similar procedures were used again in the NHES:93, as described below.

A Donnelley Marketing Information Services computer file containing census characteristics for telephone exchanges was used to stratify telephone prefixes into low and high minority concentration strata.[9] A 1992 version of the Donnelley file containing the 1990 census counts was used for the NHES:93. A current list of all prefixes in the country was obtained from AT&T and matched to the prefixes on the Donnelley file. A random sample of 50,000 telephone prefixes from the AT&T file was matched against the entire Donnelley file. The 50,000 telephone prefixes were sent to Donnelley for time and cost efficiency reasons; however, not all of these numbers were used in data collection. Any prefixes not found on the Donnelley file were assigned to the low minority concentration stratum.

We examined several different sampling plans to assess the impact of each on the sample sizes and variances of the estimates. Expected sample sizes for blacks, Hispanics and Asian/Pacific Islanders (hereafter referred to as Asians) were considered. Oversampling blacks and Hispanics reduces the numbers in the white and Asian subpopulations. In the case of whites, this loss is negligible, but in the case of Asians, it is of concern. Therefore, Asians were included in the definition of minority for the purposes of identifying those clusters to be oversampled. This procedure maintains approximately the same number of Asians in the sample as there would have been in an equal probability sample.

We concluded that an optimal design would be achieved by defining high minority concentration as at least 20 percent of either black, Hispanic, or Asian persons living in the area and sampling the high minority concentration stratum at a rate twice that of the other stratum. Thus, the high minority stratum is defined by the presence of at least 20 percent from one minority group; a prefix with 10 percent black, 6 percent Hispanic, and 5 percent Asian residency rates would not be considered a high minority stratum under this design. The design improves the precision for estimates of blacks and Hispanics without detriment to estimates for Asians and makes the overall estimates as precise as possible.

Oversampling by the characteristics of the prefix areas has two effects. First, the oversampling increases the sample sizes for minorities because they are more heavily concentrated in the prefix areas that are oversampled. Therefore, the sampling variances for estimates of these groups are reduced due to the increased sample size. On the other hand, not all of the minorities are found in the oversampled prefix areas. Thus, differential sampling rates are applied to persons depending only on their

[9] A detailed description and evaluation of such uses of the Donnelley file is given by Mohadjer (1988).

telephone prefix. As noted before, using differential rates results in increases in the sampling variances of the estimates. These increases partially offset the benefit of the larger minority sample sizes.

Some limitations of the research leading to the oversample design selection should be kept in mind when using these results.

- In estimating the percentages of households in the various oversampling strata, we assumed that the national mean household size for each racial/ethnic group applied in each stratum;
- No allowance was made for differential undercoverage in the oversampling of clusters; and
- In estimating variance ratios for the characteristics, we assumed that the population variances of the characteristics were equal within the oversampled and nonoversampled strata.

Experience with these types of approximations indicate that the overall comparisons will remain valid despite these limitations.

Data Collection Experience

The purpose of this section is to discuss the procedures used in the data collection phase of the NHES:93, including supervisor/interview staff training, data collection procedures and the use of computer assisted telephone interviewing (CATI), weekly progress in completing cases, refusal conversion, refielding other nonresponse cases, and data quality control procedures.

Supervisor and Interviewer Staff Training

A series of training sessions was held to prepare supervisors and interviewers for NHES:93 data collection. Project staff trained the supervisors and trainers on January 14, 1993. These supervisors and trainers also attended at least one full training session conducted for regular interviewers.

Interviewer training was conducted from January 15 through February 10 by supervisors and trainers in all three Telephone Research Center (TRC) locations (Gaither Road, MD; Frederick, MD; and Oceanside, CA). Based on experience with the previous NHES cycles and other similar surveys, 16 hours of training were allocated for each interviewer training group. The training program was divided into four sessions, each 4 hours in length. Immediately following training, interviewers were scheduled for a fifth 4-hour session, which was their first "live" interviewing shift on the NHES:93. This session was closely monitored by the training staff.

The goal of training was to make interviewers knowledgeable about the survey and efficient at collecting information from respondents. This included familiarizing interviewers with the questions asked in the Screener and the two extended components, the flow of the interviews, and the use of the CATI system.

Training activities included interactive and role-play scripts. Interactive scripts were delivered in lecture format, with the trainer acting as the respondent and the interviewers asking the questions. At selective points in the interactive scripts, the trainer would take time to explain or define concepts pertinent to the NHES interview to the interviewers. The role-play scripts were used to reinforce training concepts, to provide interviewers with the opportunity to practice the interview, and to provide trainers and supervisors with an opportunity to monitor the interviewers conducting a whole interview.

The first training session focused on the first interactive script, including interviews for each component. The second session focused on administration of the Screener and on avoiding refusals. Included in this session were scripts that were used to demonstrate various paths through the Screener, both for eligible and ineligible households, and a discussion of contact procedures, refusal avoidance, and the assignment of result codes. The third training session included contact procedure role plays and two additional interactive scripts, each of which was used to demonstrate multi-interview households. The fourth session was made up of six role-play interviews, in which the trainees worked at the TRC carrels and conducted interviews with one another on the telephone using prepared scripts. Supervisors monitored the floor, and training staff monitored from the TRC monitoring rooms.

Altogether six groups were trained at the Gaither Road TRC, seven groups at the Frederick TRC, and three groups at the Oceanside TRC. Generally, there were between 25 and 35 persons in each training group. In total, 450 interviewers completed training and did some interviewing for the NHES:93.

Spanish Training

Twelve interviewers were bilingual in English and Spanish. These interviewers received the same English training as all other interviewers, worked on the study conducting interviews in English for a minimum of 4 weeks, and were then trained to conduct the interviews in Spanish. One additional training session was held for these bilingual interviewers. During this additional training session, the bilingual interviewers were familiarized with the Spanish-translated version of CATI. The interviewers completed role plays and participated in interactive scripts with their bilingual trainer. The bilingual training sessions were held in Oceanside, CA and in Frederick, MD (one in each center) to accommodate the bilingual interviewers in both locations. The Spanish training session was held in the middle of March in both locations. All of the CATI screens were translated into Spanish, with the exception of contact and result screens which were not read to respondents, and the Spanish screens were available to bilingual interviewers at a keystroke.

Refusal Conversion Training

All interviewers were given strategies on how to avoid refusals during the regular training sessions. Supervisors selected experienced interviewers with higher than average cooperation rates in the Screener, the extended interviews, or both to be trained for refusal conversion. These selections were made on a flow basis as the number of refusal conversion cases available to be worked became a larger proportion of available work over time. The refusal conversion training session lasted approximately 1 1/2 hours and covered specific strategies on how to get respondents to complete an interview, common reasons for refusals, reasons specific to NHES for refusal, the importance of addressing people's concerns, and appropriate responses to respondents' concerns. The session was interactive with the interviewers sharing strategies to handle specific cases. The training sessions were first conducted about 2 weeks into data collection and were held for different groups of interviewers throughout the collection period as the work load demanded.

Data Collection Procedures

Data for the NHES:93 were collected by telephone interviewers from January 17 through May 9, 1993. While regular interviewing was closed on April 25, calls were made through May 9 to conduct data retrieval for problem cases (discussed later). Interviewers made at least seven attempts to screen households in order to determine the presence of eligible household members. These calls were staggered on different days of the week and at different times of the day over a period of at least 2 weeks, including at least two daytime calls, three evening calls, and two weekend calls. At least one weekday call was made in each of two different weeks and weekend calls were spread over at least two weekends. Cases that were coded as a problem were referred to a telephone supervisor to discuss appropriate methods of completing an interview (e.g., CATI "holding" a case in queue and releasing it for additional attempts later in the data collection period).

The CATI system scheduled cases automatically based on the algorithm that was customized for the NHES:93 survey. The system assigned cases to interviewers in the following order of priority:

- Cases that had specific appointments;
- Cases that had unspecified appointment/general callback times;
- Cases that were busy signals and came up 15 minutes later for another attempt;
- Cases that had been attempted with no contact and were tried during other specific time periods; and

- Cases that were new and had never been worked.

The NHES:93 was conducted primarily in English, but provisions were made to interview persons who spoke only Spanish. As discussed earlier, the questionnaires were translated into Spanish, Spanish versions of the CATI instruments were programmed, and bilingual interviewers were trained to complete the interview in either English or Spanish.

When the person answering the telephone was not able to speak English, and the interviewer was not bilingual and was not able to identify an English-speaking household member, the interviewer coded the case as a "language problem" and further specified the case as either "hearing/speech problem," "Spanish," or "language other than English or Spanish." Bilingual interviewers were the only ones who could access these "language problem" cases for followup. However, the bilingual interviewers were not able to access these cases until completion of the bilingual training session. Once the Spanish screens were available and bilingual training had been completed, a Spanish-speaking interviewer encountering a Spanish-speaking respondent could immediately begin to conduct the interview in Spanish without ever coding the case as a language problem[10] This occurred 33 times, and 30 of these cases were eventually completed.

The NHES:93 included a number of procedures designed to maximize the survey response rate. Since most nonresponse in an RDD survey occurs at the screening level, these procedures emphasized increasing the Screener response rate. The approaches used included refusal conversion for all Screener interviews; and refielding Screeners that had a final status of maximum calls and, in some cases, refusals. The results of these efforts are discussed later in this report.

Weekly Progress in Completing Cases

As indicated in Table 1, the percentage of interviews completed was less in the first 2 weeks of production (2 percent and 5 percent, respectively) than in comparison with later weeks (i.e., 16 percent and 18 percent in weeks five and six, respectively). This occurred for two reasons. First, interviewer training was still in progress during the first 3 weeks of production and only a portion of the interviewing staff was "live" at that time. Second, a part of the interviewing time in the early stages of an RDD survey is spent closing out nonworking and nonresidential numbers. As a result of spending time clearing these nonproductive numbers, less time is spent on the completion of cases with eligible households.

The peak of production occurred during weeks 3 through 7. By that time training had been completed, interviewers were familiar with the study, and many of the cases with nonworking or nonresidential numbers had been eliminated. The number of completed cases decreased in the last few weeks of production because the cases included labor intensive efforts such as refusal conversion, refielded nonresponse cases, and language problem cases. Each of these categories of cases has a relatively small yield compared to other cases; that is, a smaller percentage are usually completed and they frequently require a larger number of contacts per case.

The cumulative number of all completed interviews and interviewer hours is shown in Figure 1. Figure 2 shows the cumulative number of interviews by type of instrument. By March 7, about the midpoint of the data collection period, 82 percent of the total Screeners (52,053 out of 63,844) had been completed. About 77 percent of the School Readiness interviews (8,360 of 10,888), and 71 percent of School Safety and Discipline interviews (13,619 of 19,184) had been completed. As shown in both of these

[10]Prior to the availability of the Spanish CATI screens and completion of bilingual training, cases were coded as language problems by bilingual interviewers.

graphs, most interviews were completed by about April 4. The last few weeks of the data collection period were focused on the resolution of problem cases and attempts to complete cases that were refusals, had not been completed after many calls, or were language problems. As noted above, the percentage of such cases that are completed is lower than for other types of cases, and the amount of interviewer effort per completed case is greater.

Table 1.-- Percentage of interviewer hours and completed interviews, per week, NHES:93

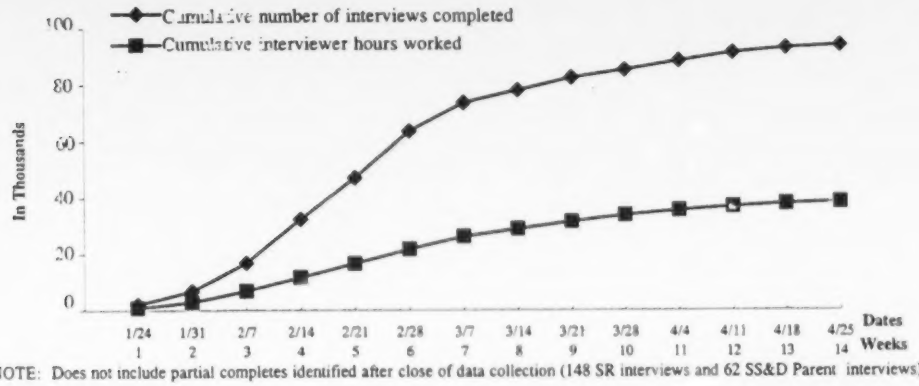
Week	Date	Percent of interviewer hours	Percent of completed interviews
1	1/24	2	2
2	1/31	5	5
3	2/7	10	11
4	2/14	13	17
5	2/21	13	16
6	2/28	13	18
7	3/7	12	11
8	3/14	7	5
9	3/21	7	5
10	3/28	6	3
11	4/4	4	3
12	4/11	3	3
13	4/18	3	2
14	4/25	2	1

NOTE: Percentages may not add to 100 percent due to rounding. Includes completed interviews at all levels (Screener, SR, and SS&D); does not include cases coded as partial completes and later kept as completes after data collection was over (148 SR parents and 63 SS&D parents).

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

In the final weeks of data collection, interviewers' hours were reduced because a lower level of daytime work was required and there were no new cases to work. The pool of cases that remained was relatively small but required more interviewer time for a comparatively smaller return. As a result of this reduction in force, there was relatively little increase in total interviewer hours during the last 4 weeks of data collection compared to the number of hours for the study as a whole.

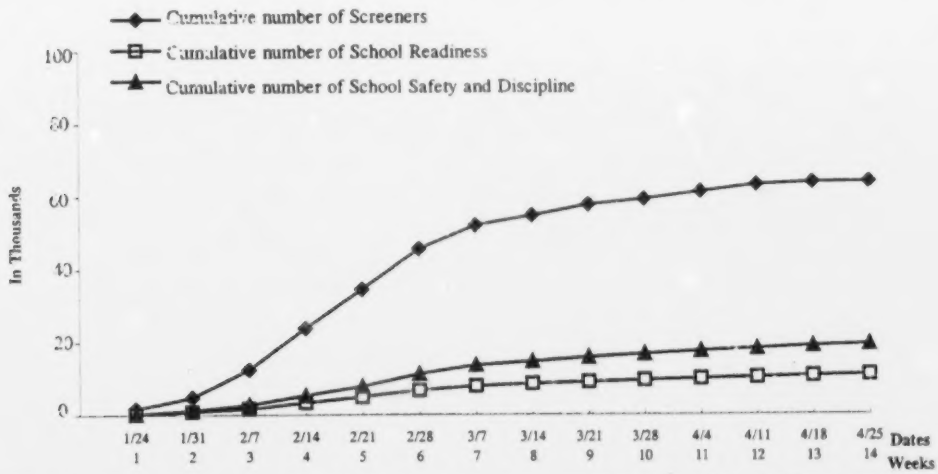
Figure 1.-- Cummulative number of interviewer hours and completed interviews, by week



NOTE: Does not include partial completes identified after close of data collection (148 SR interviews and 62 SS&D Parent interviews).

SOURCE: U.S. Deparment of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Figure 2.-- Cummulative number of completed interviews by type of interview, by week



NOTE: Does not include partial completes identified after close of data collection (148 SR interviews and 63 SS&D Parent interviews).

SOURCE: U.S. Deparment of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

BEST COPY AVAILABLE

Refusal Conversion

As a matter of standard practice, an initial refusal case is attempted again after a period of time, regardless of the type of interview (i.e., Screener or extended). If the interviewer states that the respondent was hostile or profane, the case is not refielded. (The incidence of hostile refusals is quite low; only 481 of more than 21,000 Screener refusals in the NHES:93 were ever coded as hostile; only 56 of nearly 2,000 extended interview refusals were ever coded hostile.) In general, the waiting period before a refusal conversion attempt is about 10 days to 2 weeks. Interviewers specially trained for refusal conversions are assigned to call these cases again and attempt to complete the interview. In the case of a Screener, another household member may answer the telephone and complete the Screener; in other cases, an effort must be made to convert the refusing person into a respondent.

The standard approach is to treat a case as a final refusal if the refusal conversion attempt was unsuccessful. This means that both an initial and second refusal have been obtained. In the NHES:93, an effort was made to increase the Screener response rate by refielding a subset of the "final" refusal cases. A similar effort was employed in the NHES:91 to increase response rates.

Whenever an interviewer received a refusal, information about the case was added to a CATI segment for noninterview cases. The information included a rating of the refusal as "mild," "firm," or "hostile." These ratings were, of course, subjective assessments by the interviewer. In 25 percent of the refusals, interviewers encountered a situation where the person would hang up the telephone without saying a word. In these instances, the interviewers were instructed to code the case a "mild" refusal. There is a CATI screen that gives the interviewer the opportunity to record this interaction. If a refusal was ever coded as hostile, it was not attempted again without supervisor review. Screener refusals that were coded as mild on the first attempt and either firm or mild on the second attempt, and those coded as firm on the first attempt and mild on the second attempt, were refielded for a third attempt. Those coded as firm refusals on both attempts and those ever coded as hostile were not included in the cases that were refielded.

Table 2 shows the results of the refusal conversion effort at the Screener level for the NHES:93. The overall Screener conversion rate, including initial refusals and refielding "final refusals," was 59 percent of eligible cases[11], representing 12,515 additional completed Screeners. If the cases had not been refielded, the initial refusal conversion rate would have been about 47 percent. The refusal conversion rate for the refielded Screeners was lower than the initial conversion rate, but still substantial at 35 percent (2,567 Screeners out of 7,238 total eligible Screeners).

The refusal conversion rates for the extended interviews are typically lower than the Screener rates. This is because the extended completion rates are higher and there is less room for conversion. Also, with extended interviews, the same person must be converted, while other adult household members can respond to the Screener. About 33 percent of initial refusals for the School Readiness interviews were converted, as were 32 percent of School Safety and Discipline parents, and 30 percent of the School Safety and Discipline youth (Table 3).

[11]Eligible cases include those cases that had ever been a refusal minus those later found to be nonresidential or nonworking telephone numbers.

Table 2.-- Results of Screener refusal conversion efforts on the NHES:93

Screener refusals	All refusals		Refielded refusals	
	Number	Percent	Number	Percent
Total attempted	22,327	--	7,451	--
Ineligible[1]	1,218	--	213	--
Total eligible cases	21,109	100	7,238	100
Completed	12,515	59	2,567	35
Refusal	8,185	39	4,538	63
Other nonresponse[2]	409	2	133	2

[1]Includes numbers determined to be nonhousehold numbers during conversion attempt. These cases are excluded from calculations of percents.

[2]Includes language problems, answering machines, maximum calls, and all other household nonresponses except refusals.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Table 3.-- Results of extended interview refusal conversion efforts on the NHES:93

Extended interview refusals	School Readiness		School Safety and Discipline - Parents		School Safety and Discipline - Youth	
	Number	Percent	Number	Percent	Number	Percent
Total attempted	1,336	--	1,529	--	358	--
Ineligible[1]	19	--	10	--	3[2]	--
Total eligible cases	1,317	100	1,519	100	355	100
Completed	437	33	491	32	108	30
Refusal	831	63	983	65	239	67
Other nonresponse[3]	50	4	45	3	8	2

[1]Includes cases that in the extended interview were determined ineligible due to age or enrollment.

[2]Includes cases when the youth is determined to be ineligible at the parent level or because the emancipated youth is ineligible.

[3]Includes language problems, not available during field period, sick or mentally incompetent, and duplicates of another case.

NOTE: Percentages may not sum to 100 due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Refielding Other Nonresponse Cases

As noted above, at least seven attempts were made to complete a Screener at each telephone number sampled for the NHES, with the exceptions described in the previous section. Once the number of calls was reached, the case was coded an "initial maximum call" case if a person ever answered the phone and indicated that the number was residential. The initial maximum call cases were examined by a computer program to ensure that the requisite number of attempts had been made on different days and at different times over a period of at least two weeks. In some cases, this rule was not satisfied because a household member had asked for a callback at a particular time of day or a particular day of the week. If the rule was satisfied, the case received a final status, otherwise it was refielded until it met the day and time requirements.

Another effort to increase the Screener response rate was the release of "maximum calls" cases, in which a person had answered on at least one of the seven previous attempts. The cases were held for a period of time and released for additional attempts during the last 3 weeks of the data collection period. No maximum number of calls was set for Screeners, and the cases continued to be worked until the data collection period was over. An additional 1,462 Screeners (or 20 percent of the eligible cases that were refielded) were completed as a result of refielding the maximum call cases (Table 4).

Telephone numbers that had only been answered by answering machines were classified as either residential or nonresidential based on interviewer assessments of the answering machine messages they heard. If a machine was ever coded as residential, it was classified as residential. About 19 percent (more than 24,000) of the telephone numbers in the NHES:93 sample ever had an answering machine result code, but only 1,271 were finalized as residential answering machines.

Table 4.-- Results of refielding NHES Screeners in maximum calls, NHES:93

Maximum call Screeners	Maximum calls	
	Number	Percent
Total attempted	9,023	--
Ineligible[1]	1,835	--
Total eligible cases	7,188	100
Completed	1,462	20
Refusal	376	5
Other nonresponse[2]	5,350	74

[1]Includes numbers determined to be nonhousehold numbers during conversion attempt. These cases are excluded from calculations of percents.

[2]Includes language problems, maximum calls, and all other household nonresponses except refusals.

NOTE: Percentages may not add to 100 percent due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Problems Encountered in Data Collection

Some problems were encountered during the NHES:93 data collection, as in any major collection. These generally fell into two categories. Problems related to sample management and problems related to the CATI application, which required data clarification callbacks.

Sample Management. During the early part of the data collection period, a problem in the release of secondary numbers within residential clusters was identified. Specifically, the plan had been to release 26 telephone numbers per cluster and hold back 6 telephone numbers per cluster, but all numbers were released in some early clusters. When this problem was identified, the programming was corrected for all subsequent clusters. In addition, in those clusters for which all 32 numbers had been released, but no interviews had been completed for secondary numbers in the cluster[12], the 6 telephone numbers that were erroneously released were withdrawn from active work and placed on hold as originally planned. No significant biases or inefficiencies resulted from this problem.

Data Clarification Callbacks. Five types of data clarification callback efforts were conducted. For each of these problems, the households were called back, the questions were administered, and the data were entered into the CATI data base.

1. The emancipated youths were called back to ask the household characteristics questions. These had not been collected by the CATI system when the emancipated youth completed the first interview in the household.

2. Ten emancipated youths were called back to ask if they were attending school full time.

3. One hundred ninety-three interviews had missing parent/guardian data as a result of a problem in the CATI application. These interviews were for children who had no mother or father living in the household. The households were called back to ask the appropriate questions based on the sex of the extended respondent (the mother questions were administered if the respondent was female, the father questions if the respondent was male). The remaining interviews had missing parent data due to respondent error, i.e., the mother and/or father were reported to be present in the household too late in the interview to trigger the appropriate items in the CATI. These households were called back to ask the appropriate questions for the person(s) added (mother and/or father).

4. Ten households were called back to reconcile age and grade discrepancies between two or more of the following items: S6 (AGE in Screener matrix), S9 (grade in school), S10 (grade equivalent), R1/P1 (month/year of birth), R6/P6 (grade in school), R7/P7 (grade equivalent).

5. Five interviews were readministered entirely. These were cases where the incorrect interview (School Readiness instead of School Safety and Discipline, or vice-versa) was administered due to misreporting of age and/or grade. The correct information was reported too late in the original interview to correct the survey path in CATI.

---

[12] Secondary numbers were selected and released after a primary number was identified as being residential.

Data Quality Control Activities

Several methods were used to ensure the quality of the data collected in the NHES:93. These methods included cognitive laboratory activities, CATI testing and a field test, interviewer monitoring, reinterviews, and interview taping. The procedures are described below.

Cognitive Laboratory Activities

In order to collect more reliable information in both the School Readiness and School Safety and Discipline areas, cognitive laboratory research was included in the NHES:93 design phase. This laboratory research included individual interviews and focus groups. The purpose of these activities was to provide a general evaluation of the survey questionnaires. Four areas were of particular interest: comprehension, knowledge, recall, and motivation.

The individual interviews involved two approaches: 1) debriefing a respondent following face-to-face administration of the interview, and 2) a "thinking aloud" approach in which the respondent is asked to explain what he or she is thinking about when formulating responses to the questions. Probes were used by the investigator to elicit specific information about areas of concern. For focus group activities, participants were interviewed by telephone 24 to 48 hours before the focus group and debriefing questions were explored in the focus group sessions. For additional information on cognitive laboratory activities, see the NCES technical report *Use of Cognitive Laboratories and Recorded Interviews in the National Household Education Survey* (Nolin and Chandler 1996).

CATI Testing and Field Test

Data collection quality control efforts began during the CATI development period. As the CATI system was programmed, extensive testing of the system was conducted. This testing included review by project research staff, telephone interviewing staff, data preparation staff, statistical staff, and the programmers themselves. The testing by staff members representing different aspects of the project was designed to ensure that the system was working properly from all of these perspectives. A live pretest was conducted in households between December 3 and 7, 1992; 471 Screeners, 182 School Safety and Discipline Parent, 82 School Safety and Discipline Youth, and 141 School Readiness extended interviews were completed. The purpose of this field test was to ensure that the system was working properly. Modifications to the instruments to address some administrative problems were also made at this time.

Interviewer Monitoring

Throughout data collection, supervisors and telephone monitors (experienced telephone interviewers who were trained for monitoring) monitored the interviews by listening for about 15 minutes at a time to the interviewers from either a monitoring room or from a specially-equipped carrel on the floor of the telephone center. The next major section of this report (which begins on page 24) describes the monitoring activity and results in detail.

Reinterview

A random sample of respondents who had already completed the survey was called and re-asked a subsample of items from the original interviews to check item reliability. In all, 882 reinterviews were completed for the School Readiness component and 998 were completed for the School Safety and Discipline component. The purpose of the reinterview was to:

- Identify survey items that were not reliable;
- Quantify the magnitude of the response variance for groups of items collected from the same respondent at two different times; and
- Provide feedback to improve the design of questionnaire items for future surveys.

A random sample of completed interviews was selected for reinterview. Only interviews that had never been coded a refusal were eligible. The respondent who completed the original interview was recontacted about 2 weeks after the initial interview. In order to limit the burden placed on the respondent, only a subset of items was included in the reinterview and only one reinterview per household was conducted. To avoid differential sampling of the children and youth within the various paths, the sampling within these groups was proportional to their representation in the full sample. No home schoolers or emancipated youth were sampled for the reinterview. Completed School Readiness reinterviews numbered 364 for preschool children, 163 for kindergarten children, and 355 for primary school children. Completed School Safety and Discipline reinterviews numbered 278 for parents of 6th through 12th graders, 227 for parents of 3rd through 5th graders, and 493 for youth in 6th through 12th grade.

The reinterviews for the NHES:93 were conducted using CATI, which provided an opportunity to control interviewer access to earlier responses. The entire reinterview was conducted, and then the CATI system produced a series of edit check screens to resolve differences between the initial and the reinterview responses for certain items. The check screens were used to indicate to the interviewer that two different responses had been recorded for the item, without informing the interviewer or the respondent about which response was recorded in which interview. The respondent was then asked which was the "best" answer.

Thirty-one interviewers who were currently working on the NHES:93 were trained to conduct the reinterviews. In the first week of March, all the reinterviews were done from our Frederick and Gaither Road TRC facilities. The training took approximately 3 hours. The reinterviews were conducted between March 1 and the close of data collection. Findings associated with the NHES:93 reinterview activities are included in the technical report *Reinterviews in the 1993 National Household Education Survey* (Brick et al. 1996).

Recorded Interviews

During late February and early March, 90 extended interviews were taped for later analysis. Twenty interviewers were instructed on procedures for taped interviews and were provided with tape recorders. The taped interviews are distributed evenly throughout both components.

Coders were specially trained to interpret interviewer and respondent behavior during the interview. A coding scheme was developed to assess problems the interviewer had during the administration of the interview and difficulties respondents had in terms of comprehension. Analysis of these taped interviews appears in the NHES working paper *Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey* (Brick et al. forthcoming). Findings permit the identification of problems with the questionnaires or more general problems that can be addressed through training or data collection procedures.

Monitoring Interviews in the NHES:93

Monitoring of the telephone interviewers was one of several methods used to ensure the quality of the data collected in the NHES:93. Supervisors and telephone monitors (experienced telephone interviewers who were trained for monitoring) listened to the interviewers from either a monitoring room or from a carrel on the floor of the telephone center.

Monitoring Form

The monitor was asked to complete a Monitoring Form, a two-sided form shown in appendix A, for each interview monitored. The form was used once the interviewer began regular interviewing. The first four hours of the interviewers' sessions were considered part of the training process, and the Monitoring Form was not used. Instead, the supervisors provided individualized immediate feedback to the interviewers.

The Monitoring Form was developed especially for the NHES:93 by adapting the previously used Westat monitoring form. The form covers five major areas of telephone interviewing:

- Reading and general skills;
- Listening skills and probing;
- Recording;
- Handling refusals and questions; and
- Telephone manner and relationship with respondent.

Each of these major areas is broken into subareas with two to seven specific items within each area. The Monitoring Forms were keyed and edited. If data in the specific fields did not pass the edits, the values of these items were set to missing.

The type of interview was also recorded on the Monitoring Form. The four interview types relevant to the NHES:93 are questionnaire (which will be referred to as extended for this report), Screener, refusal conversion, and language. The first two types describe the contents of the interviews that were monitored, and the last two types refer to special features of the interviews. Note that one monitoring session could cover all four types (and all four would be recorded as being included) or it could cover only one type of interview.

Monitoring Sessions

Monitors were asked to listen to the interview for approximately 15 minutes. Table 5 shows the distribution of the length of the monitoring sessions. Approximately 90 percent of all the monitoring sessions lasted between 11 and 20 minutes. About three fourths of the sessions were between 11 and 15 minutes long. Variations in time are due to interviewers completing a case and then spending time off-line completing a problem sheet, discussing a case with a supervisor, taking a break, etc.

As they listened, the monitors recorded their assessment of the interviewer's skills and abilities for each of the subitems using three categories: "no problem," "minor difficulty," "and "major difficulty." If a skill was not tested during the monitoring session, a not applicable (N/A) code was used. For example, if the session was part of an extended interview with no refusals, then one or both of the subitems in the Handling Refusals and Questions area may have been recorded as not applicable.

Table 5.-- Number of monitoring sessions, by length of session

Length of session*	Number	Percent
1 to 10 minutes	213	3.7
11 to 15 minutes	4,332	75.2
16 to 20 minutes	857	14.9
21 to 25 minutes	161	2.8
26 to 30 minutes	80	1.4
31 to 45 minutes	69	1.2
Over 45 minutes	49	0.9
Total	5,761	100.0

* Length was missing for 90 cases.

Note: Percentages may not add to 100% due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Data Reliability

Monitoring the NHES interviewers required a subjective evaluation done by persons trained for this task. The evaluation of specific interviewing skills and abilities based on monitoring is a complex and error-prone process. The interview is, in some sense, a conversation between the interviewer and the respondent. The interviewer is required to adapt the interview to the needs of the respondent without altering the interview itself. For example, the interviewer should slow down or speed up the interview, depending on the respondent's reactions to the process. No measures of the inter-rater reliability of the monitors were obtained for the monitoring. Measures of inter-rater reliability were not considered to be very important for this data collection because monitoring was not the only tool used to rate the interviewers. Without measures of the reliability, it is difficult to determine how consistently the monitors rated the interviewers. The findings presented below seem reasonable and consistent with expectations, but this does not speak directly to the reliability of the measures.

The reliability of the data is also affected by the number of sessions monitored and the number of monitors. For most regular interviewers, the number of sessions monitored was large and there were numerous monitors. However, for interviews conducted in Spanish, the number of sessions and the number of monitors was small. This is discussed further below.

Monitoring of telephone interviewers is not a procedure that is discussed much in the literature. Many evaluations of telephone interview data are mode effects studies (Groves and Kahn 1979) and do not actually examine the monitoring process. Couper, Holland, and Groves (1992) discuss sampling and monitoring procedures for a centralized telephone facility. The monitoring procedures they discuss are similar to those used in the NHES:93.

Analysis of the monitoring forms is presented in the next three sections. The first of these sections examines the monitoring reports as the survey progressed from the beginning to the end of data collection.

The next section looks in greater detail at the results by interview type. The last section summarizes the findings and suggests improvements for future studies.

Patterns Over Time

As noted earlier, the data collection period began on January 17 and ended on April 25, encompassing a period of 14 weeks. Much of the interviewing was finished early in this period, and the last weeks were spent handling refusal conversions and other problem cases as the work force was reduced. Because of the small number of interviews and monitoring sessions available at the end of the data collection period, the last two weeks are combined in the analysis that follows.

Table 6 shows the number of Monitoring Forms that were processed each week (week 13 being composed of the work done in weeks 13 and 14). As with the interviewing itself, most of the monitoring was done during weeks 3 through 7, which were the peak production weeks of the survey.

The proportion of the items that were reported as having minor or major difficulty were studied. Analysis of cases with a major difficulty showed the number of times that major difficulties were reported was too small to support meaningful findings. As a result, the minor and major difficulty categories were combined for this presentation.

Table 6.-- Number of monitoring sessions, by week

Week	Number	Percent
1 - January 17 - 23	102	1.7
2 - January 24 - 30	414	7.1
3 - January 31 - February 6	820	14.0
4 - February 7 - 13	867	14.8
5 - February 14 - 20	866	14.8
6 - February 21 - 27	684	11.7
7 - February 28 - March 6	579	9.9
8 - March 7 - 13	295	5.0
9 - March 14 - 20	381	6.5
10 - March 21 - 27	265	4.5
11 - March 28 - April 3	269	4.6
12 - April 4 - 10	216	3.7
13 - April 11 - 25	93	1.6
Total	5,851	100.0

Note: Percentages may not add to 100% due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

To examine the patterns over time, data were graphed on a control chart for proportions (P-chart). Appendix B contains one P-chart for each of the 22 items covered in the Monitoring Form. This type of control chart is commonly used in statistical quality control to examine defective rates. The chart has weeks on the x-axis and the proportion of the cases reported as having some difficulty on the y-axis. The specific

item is indicated along the y-axis using the prefix S. For example, S1A is the label for the first item in the Reading and General Skills section (1a-identifies self and reads intro clearly and without pausing).

The proportion of the records reported as having some difficulty for a given week is the plotted point. The mean proportion across all 13 weeks is called *P* and is shown as a center line on the chart. Each chart also has an upper control limit (UCL) and lower control limit (LCL) which are three standard errors above and below the average proportion, respectively. Since the sample size varies from week to week, the UCL and LCL also vary. Points above or below the control limit indicate something in the system has changed from a steady process (one in statistical control) and would be worth investigating.

The P-chart is a useful device for graphically displaying the data from the monitoring sessions. The purpose of displaying the data in this fashion is to highlight the important general patterns and identify those data that depart from this pattern. It does not provide immediate feedback, which is an important function of modern statistical quality control. However, the P-charts do provide a useful means of viewing an activity as a whole.

General Patterns

Below we discuss some of the general patterns that are observed in the charts (see appendix B, figures 1 through 22) and probable reasons for those patterns. Although these general patterns can be seen in most of the charts, figure 1, for item S1A, identifies self and reads introduction clearly (page B-1), is a good illustrative example for many of these patterns.

The overall proportion reported as having some difficulty was generally small. The value of *P* was 0.05 or less for all but 6 of the 22 items (see appendix B for all control charts). The maximum value of *P* was 0.085.

The weekly proportion was typically largest (often exceeding the UCL) during the early weeks of data collection, typically around the second to the fourth week. This pattern is consistent with both a learning curve for a new survey and the introduction of new telephone interviewers to the workforce. As with any new survey, interviewers became more familiar with the wording and concepts as time progressed and the number of cases with some difficulty decreased with this experience. Because experienced interviewers were trained first, the curve typically was higher in the second or third week when the experienced interviewers were joined in the workforce by the newly trained interviewers.

After the early peaks, the weekly proportions generally dropped dramatically. This was the result of both the learning curve and the elimination of interviewers with poorer than average performance. Interviewers with poor quality performance were released after they were given a chance to improve their skills. Later, as the workload decreased, only interviewers with high performance ratings were retained.

In several cases, the weekly proportion increased again in the last weeks of the period, although not usually to the peak levels. This type of behavior is visible in appendix B, figures 2, 20, and 21. A greater increase in the last week of data collection is observed in figure 10. This rise was probably due to the type of work that remained during the last weeks, almost all of which involved either refusal conversions or language problems. In some of these situations, the interviewers may have deviated more substantially from the prescribed protocol to gain the cooperation of the respondent. We will come back to this point in the next section when the results are analyzed by type of interview. It is also worth noting that the number of cases was small in the last weeks of data collection. As a result, the estimated proportions are subject to more variability.

Another common pattern was one in which the estimated weekly proportion was relatively flat over the entire period. Figures 5, 7, 12, 19, and 22 in appendix B have this shape. These measures include verifying spelling of names, addresses, and so on; listening to the entire answer; using comments appropriately, offering the 800 number for verification; and avoiding personal involvement. Apparently, these aspects of interviewer behavior were less subject to change over time as a result of accumulated experience.

In many aspects, the general patterns exhibited in the charts are consistent with expectations. During the initial periods of learning and experiencing the interview process, the monitors noted more difficulties. Later, these difficulties became less common and dropped significantly until the mix of work had changed significantly at the very end of the data collection period. While this pattern existed for most items, some showed little or no change over the entire period.

Interview Type

In this section, the percentage of cases reported as having some difficulty is presented for all of the items in each of the five major areas. The percentage with difficulty is also presented by the interview type: extended interview, Screener interview, refusal conversion interview, and language problem interview. As noted before, one monitoring session might span multiple interview types.

The estimated percentage of the cases in which the interviewer had some difficulty with a skill is presented for each major area by interview type. Table 7 shows the results for the Reading and General Skills section of the Monitoring Form. Tables 8 through 11 contain the estimates for the other four major areas. All five tables have the same structure and format.

The first row of each table gives the overall percentage of the cases reported as having some difficulty. Associated with each estimated percentage is the base for the percentage, which is labeled as the number eligible. The number eligible is less than the number in the total column because not every monitoring session resulted in an evaluation from the monitor for each item and because the data items that failed the edit were set to missing.

Overall Rates

The estimated percentage in the total row of each of the five tables shows that the percentage of cases where the interviewer had some difficulty with a skill was relatively small, as noted earlier. The item with the largest estimated percentage was the second item in the Reading and General Skills area; 8.5 percent of the cases were reported as having some difficulty in reading all appropriate phrases and answer categories (1b).

The major area with the highest reported percentages was Handling Refusals and Questions (see table 10). The percentage with difficulty answering respondent questions and objections clearly, confidently, and briefly without hesitation (4a) was estimated at 8.4 percent. The percentage who had difficulty with offering the verification number (4b) (so the respondent could call and verify the validity of the survey) was 6.7 percent. Extensive training in this area was conducted for the NHES:93 because this is a common problem in RDD surveys.

Extended Interview and Screener Rates

The next four rows of the tables give the estimated percentage with difficulty for the monitored session by whether or not an extended interview or a Screener interview was done in the session. Roughly 80 percent of all sessions had an extended interview and 70 percent of all sessions had a Screener interview.

The estimated percentage with difficulty was relatively consistent between interviewing sessions with and without screeners and with and without extended interviews for the vast majority of the 22 items. The percentages with difficulty generally did not differ by more than 1 percent across the items.

Some of the items that exhibited significant variation for extended and Screener interviews measured skills that were predominately only tested in one of the two types of interviews. These items include verification of spelling (1e), remaining neutral (2d), recording results (3e), moving the matrix (3f), and refraining from opinions (5b). For example, verification of spelling was generally only requested in the Screener when names of household members were requested. Remaining neutral or refraining from opinions was relatively simple during the Screener, because it did not involve questionnaire items associated with positive or negative value judgments.

The greatest differences across extended and Screener interview types were in the Handling Refusals and Questions area. The percentage with difficulty answering respondent questions (4a) was an estimated 12.7 percent for sessions where there was no extended interview. This estimate combined with the 8.9 percent estimated for sessions with Screeners and the 8.6 percent estimated for the verification item (4b) for sessions with no extended interviews, form a pattern. *Together, these results indicate that the problems noted by the monitors were typically handling respondent queries in the initial stages of the screening interview.* This skill is the most critical to gaining the cooperation of the respondent to complete the rest of interviews.

Refusal and Language Problem Rates

The last four rows of the tables give the estimated percentage for sessions by whether or not the interviewer was doing refusal conversion or language problem interview during the session. About 15 percent of the monitored session included at least some refusal conversions, while only about 2 percent of the monitored sessions were language problem cases.

Since there were only 100 language problem cases among the monitoring forms, estimates for language problem interviews are subject to higher sampling errors than the other types. The impact of one particular interviewer may also be more dominant when the sample size is small. For example, with only a few Spanish language interviewers for the NHES:93, it is possible that one interviewer having greater difficulty could have been monitored multiple times and have a substantial impact on the estimates. In addition, the number of monitors for language problem interviewers was smaller, generally only two persons at each telephone center. Because ratings given in monitoring are subjective, the small number of monitors could have an impact on the estimates. These factors should be considered in interpreting the results below.

For both refusal conversion and language problem cases the estimated percentage with difficulty varies quite significantly from item to item across the major areas. In many respects, this reflects the differences that are inherent in the type of work. For example, in refusal conversion cases, the interviewer may intentionally avoid verifying the spelling so the respondent will not have a chance to interrupt and break off the interview.

In all but two of the items, the cases that involved refusal conversions had a lower estimate of the percentage with difficulty than the cases that did not involve refusal conversion. The two exceptions were items related to the pace of the interview (1b and 1e), which the refusal converters may have intentionally decided to relax for this type of work. The explanation for the higher ratings for the other items is related to the characteristics of the interviewers who did refusal conversions. Interviewers were trained for refusal conversion if their performance on the regular work was above average. Thus, findings of lower problems with the skills of the refusal conversion work is very understandable.

The language problem sessions were subject to substantial variability across the skills. For the items from the Recording Skills and the Handling Refusal Skills area, the language problem cases had somewhat smaller estimates of the percentage with difficulties than for the sessions that did not include language problems. In most of the items in the other areas, this was also true.

One of the items that showed significant variability for language problem sessions was reading all appropriate phrases (1b), where the percentage with some difficulty was nearly 15 percent. This skill may have been related to the *quality of the Spanish translation* of the instrument and the fact that there are several different dialects of Spanish. The Spanish interview was not as polished as the English version, and interviewers were required to improvise in some situations. This problem was exacerbated by respondents who spoke different dialects of Spanish. These problems may have encouraged the language problem interviewers to stray from the text for questions that should have been asked exactly as written.

Two items in Listening and Probing Skills area had results which went in opposite directions. *The percentage with some difficulty listening to entire answer (2a) was higher than the overall average, while the percentage with difficulty listening for what may not be said and probing (2b) was much lower than average.* Some of these differences may be artifacts of the small sample size.

The biggest problems for language problem sessions were the higher percentages estimated to have difficulty with remaining neutral (2d) and refraining from opinions (5b). For both items, the percentage with difficulty was much larger for the language problem cases than for the other cases. It is not clear from these results if these estimates are due to consistent problems across interviewers, or are associated with one or two specific language problem interviewers.

Table 7.--Summary of reading and general skills from monitoring forms, by type of interview

Interview characteristic	Total	Identifies (1a)		Reads phrases (1b)		Follows skips (1c)		Reads clearly (1d)		Verifies spelling (1e)		Adjusts pace (1f)	
		Number Eligible	Percent Difficulty	Number Eligible	Percent Difficulty	Number Eligible	Percent Difficulty	Number Eligible	Percent Difficulty	Number Eligible	Percent Difficulty	Number Eligible	Percent Difficulty
Total	5,851	4,004	4.5%	4,896	8.5%	880	0.9%	5,134	2.0%	2,600	5.0%	3,533	3.1%
Extended yes	4,378	2,711	3.8%	3,836	8.4%	655	0.8%	3,989	2.1%	2,054	5.5%	2,810	3.1%
Extended no	1,163	1,028	4.2%	828	7.0%	174	1.1%	895	1.7%	416	2.9%	536	3.7%
Screener yes	3,869	3,097	4.4%	3,187	8.5%	544	1.1%	3,375	1.7%	1,740	5.1%	2,271	3.0%
Screener no	1,681	736	3.5%	1,454	7.8%	277	0.7%	1,490	2.3%	724	4.7%	1,063	3.2%
Refusal yes	748	617	1.8%	595	6.9%	74	0.0%	637	0.6%	298	7.0%	324	0.9%
Refusal no	4,474	2,958	3.9%	3,812	8.3%	699	0.6%	3,958	2.0%	2,024	4.5%	2,791	3.2%
Language yes	100	61	3.3%	88	14.8%	37	0.0%	90	2.2%	61	6.6%	73	2.7%
Language no	5,166	3,543	3.5%	4,357	8.0%	734	0.5%	4,542	1.8%	2,280	4.7%	3,060	3.0%

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Table 8.--Summary of listening and probing skills from monitoring forms, by type of interview

Interview characteristic	Total	Listens to entire (2a)		Listens for not (2b)		Probes unclear (2c)		Remains neutral (2d)	
		Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty
Total	5,851	4,755	1.9%	3,371	4.7%	3,693	5.3%	3,637	3.7%
Extended yes	4,378	3,840	1.8%	2,876	4.9%	3,214	5.5%	3,138	4.0%
Extended no	1,163	711	1.5%	386	3.4%	362	3.9%	381	1.3%
Screener yes	3,869	3,033	1.7%	2,028	4.0%	2,180	4.6%	2,170	2.8%
Screener no	1,681	1,468	2.0%	1,141	5.5%	1,310	6.3%	1,261	4.8%
Refusal yes	748	575	0.7%	351	2.8%	392	3.1%	392	1.3%
Refusal no	4,474	3,678	2.0%	2,682	4.6%	2,926	5.7%	2,876	3.9%
Language yes	100	87	3.4%	85	0.0%	86	1.2%	85	15.3%
Language no	5,166	4,202	1.7%	2,965	4.6%	3,260	5.5%	3,209	3.2%

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Table 9.--Summary of recording skills from monitoring forms, by type of interview

Interview characteristic	Total	Records accur. (3a)		Uses comments (3b)		Corrects errors (3c)		Uses CTRL keys (3d)		Records results (3e)		Moves matrix (3f)		Uses HH select (3g)	
		Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty
Total	5,851	5,056	5.2%	999	5.0%	2,837	2.5%	2,272	2.3%	3,911	5.5%	1,533	4.0%	2,178	2.3%
Extended yes	4,378	3,980	5.1%	772	5.1%	2,388	2.5%	1,824	2.1%	2,686	4.8%	1,177	3.0%	1,898	2.1%
Extended no	1,163	842	4.8%	201	4.5%	355	2.0%	366	2.2%	987	5.9%	273	4.4%	216	2.2%
Screener yes	3,869	3,296	5.5%	601	5.3%	1,694	2.6%	1,368	2.6%	2,954	6.3%	1,226	4.6%	1,397	2.6%
Screener no	1,681	1,500	4.1%	357	4.8%	1,002	2.2%	795	1.6%	821	2.7%	250	0.0%	664	1.6%
Refusal yes	748	638	2.0%	156	4.5%	338	0.6%	266	1.1%	607	3.1%	179	1.1%	213	1.1%
Refusal no	4,474	3,894	5.3%	775	4.9%	2,185	2.7%	1,805	2.3%	2,918	5.4%	1,172	4.0%	1,711	2.3%
Language yes	100	88	3.4%	36	2.8%	49	0.0%	81	0.0%	55	0.0%	32	0.0%	49	0.0%
Language no	5,166	4,483	4.9%	894	4.9%	2,479	2.5%	2,015	2.2%	3,497	5.1%	1,330	3.7%	1,891	2.2%

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Table 10.--Summary of handling refusals skills from monitoring forms, by type of interview

Interview Characteristic		Total	Answers questions (4a)		Offers verification (4b)	
			Number eligible	Percent difficulty	Number eligible	Percent difficulty
Total		5,851	2,523	8.4%	297	6.7%
Extended	yes	1,163	528	12.7%	81	8.6%
Extended	no	4,378	1,865	6.6%	205	4.4%
Screener	yes	1,681	665	6.0%	55	5.5%
Screener	no	3,869	1,772	8.9%	224	6.3%
Refusal	yes	4,474	1,824	8.8%	183	6.6%
Refusal	no	748	442	3.6%	85	2.4%
Language	yes	5,166	2,241	7.9%	254	5.5%
Language	no	100	46	4.3%	14	0.0%

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

Table 11.--Summary of telephone manner skills from monitoring forms, by type of interview

Interview characteristic	Total	Is pleasant (5a)		Defines opinions (5b)		Accepts emotion (5c)	
		Number eligible	Percent difficulty	Number eligible	Percent difficulty	Number eligible	Percent difficulty
Total	5,851	5,239	2.4%	4,484	4.7%	2,557	1.2%
Extended yes	4,378	4,032	2.5%	3,615	5.0%	1,931	1.2%
Extended no	1,163	948	2.2%	673	2.2%	476	0.8%
Screener yes	3,869	3,477	2.2%	2,940	3.7%	1,740	0.9%
Screener no	1,681	1,493	2.6%	1,318	6.1%	670	1.5%
Refusal yes	748	668	1.5%	586	1.7%	297	0.3%
Refusal no	4,474	4,014	2.5%	3,447	4.8%	1,976	1.2%
Language yes	100	94	3.2%	82	11.0%	68	1.5%
Language no	5,166	4629	2.4%	3,979	4.2%	2,226	1.1%

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

BEST COPY AVAILABLE

Summary

The Monitoring Form was used in the NHES:93 to obtain some measure of the performance of interviewers in five major skill areas. Although the Monitoring Form results provide subjective evaluations of the interviewers, it is important to realize that other aspects such as the CATI instrument, the items on the questionnaires, the language of the questionnaires and the language of the respondents are part of interview success. See a separate NHES:93 working paper, *Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey* (Brick et al. forthcoming), for an analysis of a sample of tape recorded interviews and the impact of the instruments on the quality of the data from the survey.

The percentage of sessions in which the interviewers were reported to have some difficulty with any particular skill was generally small. Reading all appropriate phrases and answer categories was the item with the largest estimated percentage having some difficulty. The interviewers are trained and expected to read the items exactly as written, but many deviate from the exact wording on occasion. The impact of these deviations on the quality of the data is not completely clear. Some (see Suchman and Jordan, 1990) suggest that less structure makes the interview more conversational and improves the process. While this is not encouraged, minor deviations are not serious in most cases. This problem is evaluated more fully in the working paper described above.

Handling Refusals and Questions was one area that had consistently higher than average reported difficulty. This finding is consistent with the importance placed on handling refusals and questions in the training of interviewers. The reported problems associated with this area showed a marked decrease after the first several weeks of interviewing, when the interviewers had some experience in the interview. These results suggest that handling refusals and questions needs to remain an important topic in training, and that investigating new approaches to this training might be profitable.

The pattern of having a peak of difficulties relatively early in the interview process followed by a marked decline is consistent with the expected learning curve and the introduction of new interviewers. As interviewers became more familiar with the questions, the number of cases with some difficulty decreased. The improvement in ratings was further enhanced by the selective retention of interviewers with higher than average performance ratings. This type of pattern is not likely to change for the NHES because of the need to complete a large number of interviews in a short period of time.

Because some of the skills measured on the Monitoring Form were exercised more often in one of these two different types of interviews, some differences in results were expected. However, the observed differences are quite minor; the estimated percentage with difficulty did not vary in important ways for extended and Screener interviews. The results for the two interview types also highlighted the importance of answering questions to gain the cooperation of the respondent. Once cooperation was obtained, the interviews proceeded relatively well.

Refusal conversion sessions (748 monitoring sessions) had lower percentages of cases with difficulty than the monitored sessions that did not involve refusal conversion. This result was largely a function of the process of selecting interviewers to be refusal converters. The percentage of cases with difficulty was higher for the refusal conversion interviewers for two items related to the pace of the interview. It is very possible that deviations from the normal pace may have been a reaction to the respondents and actually improved response rates.

The language problem sessions were subject to substantial variability across the skills. For many items, the language problem cases had lower estimates of difficulties than for the sessions that did not include language problems. However, language problem cases had a very high percentage of reported difficulties with reading all appropriate phrases. It is very likely that this skill may have been related to the quality of the Spanish translation of the instrument. Although some of these problems may be related to different Spanish dialects used among Hispanic persons in the U.S., future surveys should invest more time and resources into improving the quality of the translated instruments.

Another problem noted for language problem cases was difficulty with remaining neutral and refraining from opinions. Since there were only 100 language problem sessions monitored, it is not clear if these estimates are due to real differences or the fact that one or two interviewers or monitors might have had a substantial impact on the estimates. Future studies of language problem cases will involve monitoring more than 100 sessions to help clarify the results.

Survey Administration Time for the NHES:93

The purpose of this section is to report the response burden for the NHES:93 Screener, School Readiness, and School Safety and Discipline interviews. The time it takes a respondent to complete a survey interview is an important factor in both the response rate and response quality. While surveys need to include all of the important analytic variables, they should also strive to be as brief as possible to reduce the burden on the public and encourage complete and reliable responses. For the NHES:93, the amount of time it took to complete various components of each of the three major interviews was automatically recorded on the database. These administration times can be used to measure response burden, to measure the productivity of interviewers, and to plan for future studies using similar items.

The next section describes the procedures used to edit the interview timing data collection in the NHES:93. The edited data are used in the subsequent sections to describe the experiences over the entire data collection period and for particular types of interviews.

Editing the Administration Time

The time required to complete each segment of each interview was recorded automatically by the NHES:93 CATI system. However, this recorded time does not always reflect the true administration time. For example, if the interviewer waited on the telephone while the respondent took care of other business, such as answering the door or tending to a child, the length of time recorded would be artificially inflated. In these instances, the interviewer had no formal way to record why the interview was taking longer than normal. Monitoring of interviews allowed project staff to be able to see that such delays in interviews do occur and that provisions need to be made to give an accurate representation of the administration length.

Because the purpose of this analysis is to estimate respondent burden for the NHES:93 interviews, recorded times that were extreme outliers were edited. The process of editing the outliers involved analyzing the frequency of administration times in each of the 24 segments making up the three interviews (1 Screener segment, 14 School Readiness (SR) segments, 9 School Safety and Discipline (SS&D) segments). The mean time was assigned to the top and bottom 1 percent of outlying scores for 23 of the 24 segments. The remaining segment (the SR segment on television viewing) was assigned the mean value to cases at the top 1 percent and bottom 5 percent of cases. The procedure for this segment differed from the others because the recorded administration times in this segment showed more variability in the outlying scores. These cutoffs were determined after examination of the frequency distribution of administration times. The procedures used for the outliers are different from those used in the NHES:91 because the distribution of timings in the two years were different.

Two other situations required special treatment. Timing data for the emancipated youth path in the SS&D interview were incomplete due to a programming problem. The emancipated youth were not asked the final segment (questions P116 through P122) during the CATI interview, and these 77 cases were called back and administered a pencil and paper interview to obtain these data. Therefore, the time it took to administer this segment was not recorded in the database for these interviews. In addition, the timing variable was not set for the first segment (questions P1 through P9a) for the emancipated youth path, so the timing data were not recorded for those interviews. For cases that had missing timing values, the mean segment time was imputed from interviews with respondents other than emancipated youth so as not to underestimate the total interview time.

Table 12 at the end of this section presents the administration times in minutes for the Screener, the SR component, and the SS&D component. The times are presented for interviews that are considered complete; i.e., the respondent provided answers to all items considered critical to fulfilling the purpose of the survey. The total interview times, in addition to the high item-response rates obtained, indicate that the NHES:93 was relatively successful in obtaining the required data without overburdening the respondents. These overall averages, however, are less informative than the timings for the various paths (e.g., the SR preschool path or the SS&D youth path) because of the extensive skip patterns within each interview component. These are discussed in the sections that follow.

Screener Administration Time

The actual time needed to administer the Screener, as shown in table 12, was somewhat less than anticipated prior to conducting the survey. The overall mean time was 2.5 minutes, which is about 1.5 minutes less than the time estimated in the OMB package. This mean time was subject to substantial variability; the standard deviation was 1.5 minutes, and half of the interviews took between 1.5 and 3.3 minutes to administer.

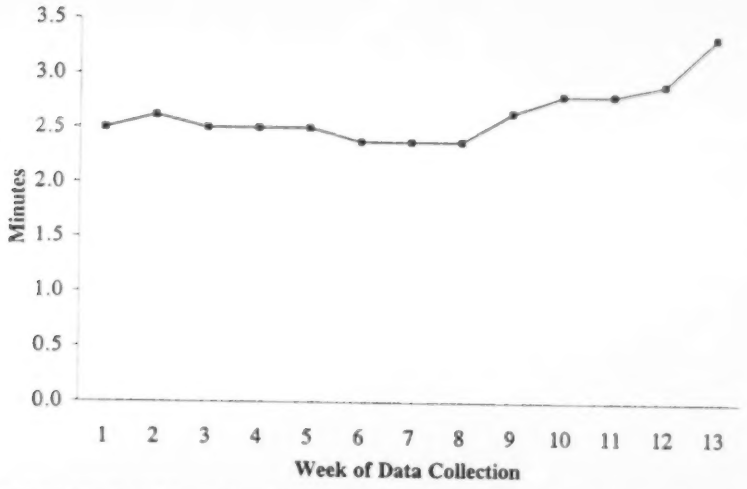
The time required to administer the Screener varied because the number of questions asked in each household depended on the household composition. If no one in the household was 18 years old or younger or in the 12th grade or lower, then the Screener was complete without enumerating the household members. On average, this path (ineligible due to age and enrollment) took only 1.7 minutes to complete and had very little variability.

For all other paths, the members of the household were enumerated. The mean for the shortest of these paths was 3.2 minutes; this was the path for households that were enumerated but had no one eligible for the sample after the enumeration was completed. For example, if all children were under age 3, information about the child's enrollment status, grade, and most knowledgeable parent was not collected. The mean time to complete the Screener was slightly longer for households that had an eligible child who was not sampled (3.6 minutes). In households with children sampled only for the SS&D component, the mean time was 3.9 minutes. For households with children sampled only for the SR component, the mean time was virtually the same (3.8 minutes). In general, enumeration took longer when there was more than one child in the household. The Screener took the longest when children were sampled for both the SS&D and the SR extended interviews. The mean for these households was 4.9 minutes. The increase in the time for these households was generally due to the households having more members to enumerate, and the identification of the parent or guardian respondents for the extended interviews.

Effect of Week of Data Collection on Screener Administration Time

Figure 3 presents the administration times of the Screener for each week of the data collection period. During the first 8 weeks of data collection, the Screener took about 2.5 minutes to administer (with the weekly means between 2.4 and 2.6 minutes). In the initial weeks of the survey, several different things happened that affected performance (including training of interviewers and releasing interviewers with poor performance). During these initial weeks, the mean administration time for the Screener was relatively insensitive to these activities.

Figure 3.--Mean administration time in minutes* of Screener over data collection period



*Does not include partial completes identified after close of data collection (148 SR interviews and 63 SS&D Parent interviews).
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

During the final 5 weeks of data collection, the time of administration consistently increased. The mean time for Screeners conducted in the 13th week was 3.3 minutes. The increase in administration time during this phase occurred for two reasons. First, toward the end of data collection, more problem cases were called, sometimes involving additional questions or clarifications. Second, much of the work at the end of the data collection period involved time-consuming language problem cases and refusal conversions. This result is typical of most RDD surveys. The NHES:91 did not follow this pattern. Rather, in that survey, the screening time declined during the later part of data collection because sampling for the Adult Education component was terminated during the survey period while the other component, Early Childhood Education, was still being conducted. This had the effect of reducing the number of households in which any enumeration was necessary.

School Readiness Administration Time

The mean time to administer the SR interview was 21.5 minutes, about 3 minutes longer than expected when the survey was planned. Half the SR interviews took between 18.1 and 24.5 minutes to complete. The mean time to complete the interview was nearly the same for all three major paths (Preschool, Kindergarten, and Primary). Only the Home School path had a substantially lower mean time to complete. Its mean time was about 6 minutes less than the other paths, mainly because entire segments of the interview were not included for these children.

When more than one child was sampled from the household, some data items were collected only once per household or once per respondent. This affected the administration time recorded for the parent information and household characteristics segments of both the SR and SS&D interviews. For example, household characteristics like income and ZIP code were asked only in the first extended interview conducted in the household. Other items, like mother's education, were asked only one time if the same person was the mother of more than one sampled child in the household. The rationale for collecting the information one time for all sampled members is obvious. However, this timesaving device does complicate the analysis of the timing data, since the time is recorded only for the first extended interview in the household. The first interview in a household was usually an SR interview, but also could have been an SS&D interview.

The impact of this method of collecting the data is to slightly suppress the mean time to complete each extended interview while correctly reporting the overall response time for extended interviews. When the mean time to complete a particular extended interview is discussed, this factor should be kept in mind.

Table 12 shows the mean times to collect various segments of the SR interview. For example, the introduction consisted of questionnaire items R1 through R8. These items were asked of all completed interviews and had a mean time to complete of just under 1 minute. Other segments were not asked for some of the paths. For example, the developmental profile items were asked only for preschool children. These mean timings can be useful in evaluating the utility of the items for future studies.

Four of the segments took more than 2 minutes to complete on average: television viewing (3.6 minutes), health and nutrition (2.8 minutes), developmental profile (2.6 minutes), and teacher feedback (2.3 minutes). Background items on the child, household, and parents were collected in the introduction, the parent information segment, and the household characteristics segment. When these three segments are combined, the mean time to administer the background items was about 4 minutes.

School Safety and Discipline Administration Time

Prior to the start of data collection, the mean time to administer the SS&D interview was estimated to be about 18 minutes per interview. The actual administration time across all SS&D interviews was just over 16 minutes.

The mean administration time varied substantially by the type of interview and the major path. The interviews with parents of 6th through 12th graders were the longest, taking an average of 19.3 minutes to complete. This was nearly 5 minutes longer than the interviews with parents of 3rd through 5th graders and the interviews with emancipated youth. The interview with youth took only 12 minutes to complete. This interview was shorter than the parent interviews because it involved only a subset of items asked of the parents and only 28 additional items not asked in the parent path. As noted earlier, the emancipated youth timings include imputed times for the introduction and the household characteristics segments.

The last section of table 12 gives the mean times to complete various segments of the SS&D interview. These mean times are aggregates across the three different paths and exhibit some variability as a result. For example, some school environment items were asked in all of the paths, but the number of items varied by the type of respondent. For some segments, like the one on school discipline policy, this variation is evident.

It should be noted that for the SR and the SS&D interviews, the introduction, the parent characteristics, and the household characteristics segments are parallel. The mean times for these segments reflect this similarity.

Table 12.-- Mean, median, and quartile administration time (in minutes) of NHES:93 completed interviews, by interview type and segment

Completed interviews and interview components	Number	Interview length in minutes				
		Mean	Standard deviation	Quartiles		
				75%	Median	25%
Interview totals:						
Screener	63,844	2.5	1.5	3.3	2.0	1.5
School Readiness	10,888	21.5	5.0	24.5	21.5	18.1
School Safety and Discipline	19,184	16.1	4.9	19.2	15.8	12.2
Interview type:						
Screener						
Eligible, but none sampled	802	3.6	1.2	3.9	3.3	2.7
Both School Readiness and School Safety and Discipline sampled	3,432	4.9	1.6	5.6	4.5	3.8
Only School Safety and Discipline sampled	9,392	3.9	1.2	4.5	3.7	3.1
Only School Readiness sampled	6,594	3.8	1.1	4.3	3.6	3.0
Ineligible due to age and enrollment	40,972	1.7	0.5	1.9	1.6	1.4
Ineligible, but enumerated	2,652	3.2	1.0	3.6	2.9	2.5
School Readiness						
Preschool	4,423	22.0	4.4	24.5	22.0	18.9
Kindergarten	2,126	22.6	5.1	25.5	22.6	19.1
Primary	4,277	21.3	4.8	24.1	21.2	18.0
Home School	62	15.7	4.2	17.6	15.4	13.6
School Safety and Discipline						
Parents of 6th-12th graders	10,117	19.3	4.1	21.6	19.2	16.4
Parents of 3rd-5th graders	2,563	14.7	3.6	16.9	14.6	12.2
Youth in 6th-12th grade	6,427	12.0	2.4	13.1	11.7	10.4
Emancipated youth*	77	14.7	3.3	16.3	14.6	12.6
Interview segments:						
School Readiness						
Introduction (R1-R8)	10,888	0.8	0.4	0.9	0.6	0.5
Family Relationships (R10-R13)	10,888	0.8	0.4	0.9	0.7	0.5
Developmental Profile (R14-R31)	4,423	2.6	0.7	2.9	2.4	2.1
Program Participation (R32-R49)	10,888	1.6	1.1	2.3	1.4	0.7
School Adjustment (R51)	6,403	1.2	0.4	1.3	1.0	0.9
Teacher Feedback (R52-R53)	6,403	2.3	0.5	2.3	2.3	2.1
Kindergarten Items (R55-R70)	10,696	1.4	0.7	1.7	1.3	0.9
Primary Items (R71-R85)	6,403	1.3	0.9	1.8	1.4	0.3
Home Activity (R86-R91)	10,888	1.0	0.3	1.1	0.9	0.8
Television Viewing (R92-R100)	10,888	3.6	1.4	4.6	3.6	2.2
Health/Nutrition (R101-R118)	10,888	2.8	0.9	3.3	2.6	2.0
Experiences < Age 5 (R119-R131)	6,403	1.9	0.7	2.2	1.7	1.4
Parent Information (R132-R157)	10,888	1.5	1.0	2.0	1.5	0.7
Household Characteristics (R158-R168)	10,888	1.7	0.4	1.8	1.6	1.4

Table 12.-- Mean, median, and quartile administration time (in minutes) of NHES:93 completed interviews, by interview type and segment--Continued

Completed interviews and interview components	Number	Interview length in minutes				
		Mean	Standard deviation	Quartiles		
				75%	Median	25%
School Safety and Discipline-Parents						
Introduction (P1-P9A)	12,680	1.0	0.6	1.2	0.9	0.7
School Characteristics (P10-P20)	12,680	1.9	0.6	2.1	1.7	1.5
School Environment (PY21-P25)	19,184	2.1	0.7	2.4	2.0	1.7
School Safety (PY26-PY55)	19,184	1.7	0.7	2.0	1.6	1.2
School Discipline Policy (P56-Y60)	19,184	0.6	0.4	0.9	0.5	0.3
Substance Use and Education (Y61-Y69)	19,184	2.1	0.8	2.6	2.1	1.6
Child Characteristics (P70-Y98)	19,184	4.3	1.8	5.5	4.3	2.6
Parent Information (P99-P114)	12,243	1.1	0.5	1.3	1.0	0.7
Household Characteristics (P116-P123).	12,680	1.5	0.4	1.7	1.5	1.3

*The time for these interviews included imputed times for the introduction segment (P1-P9A) and the household characteristics segment (P116-P123).

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey (NHES), spring 1993.

NHES:93 Data Editing

The final product of the NHES CATI data collection process is the delivery of edited data files and associated documentation. The ultimate usefulness of the data collection will be determined by the quality of these products. In order to ensure data quality, a series of post-data collection data editing procedures were conducted. Data editing (correcting interviewer, respondent, and program errors) was performed throughout the NHES data collection and potentially introduced other errors in data items that had previously been edited during the CATI administration. In addition to data editing, data retrieval efforts required entry of data outside the CATI program. Therefore, extensive post-data collection data editing procedures were instituted to achieve quality data. These editing procedures included checking data alignment, confirming that data were within the defined range of values for each data item, performing consistency and structural edits between data items, reviewing cross-tabulations between data items, reviewing frequency distributions for individual data items to ensure skip patterns were followed appropriately, and reviewing text responses of "Other, Specify" fields for coding. After imputation of missing values was completed, these procedures were repeated to ensure that no errors were introduced during imputation.

Data Alignment

Character data are left justified ("John "). Numeric data are right justified (" 200.25"). After data collection, alignment edits were run against the entire database to ensure appropriate alignment of data. This provided for clean frequency review by representing all identical values together. For example, " 1" and "1 " are represented in the database as " 1".

Range Edits

The ranges of most items in the NHES CATI surveys were determined by the codes available for the responses, since they were close ended. For open-ended items requiring an entry by the interviewer, such as ages, dates, and number of nursery schools child attends, there was not a specific set of responses. Therefore, reasonable ranges were defined and applied to these items.

Range checks included both hard and soft edits. A soft range is one that represents the reasonable expected range of values, but does not include all possible values. Responses outside the soft range were confirmed with the respondent (i.e. the answer was repeated to the respondent for confirmation or correction) and reentered. For example, the age a child was when he or she first attended any Head Start program had a soft range of 3 to 5. A value outside this range could be entered and confirmed as correct as long as it was within the hard range of values (2 to current age). Hard ranges are those that have a finite set of parameters for the values that can be entered into the CATI system. The NHES CATI system did not accept out-of-hard-range values for either open- or close-ended questions when an acceptable range of values was defined. Such values were entered in the CATI comments utility, after confirming with the respondent that they were correct, and the variables were updated during the data preparation process.

After data collection was completed, range edits were rerun against the entire database to ensure that no outliers (other than those specified in interviewer comments as being legitimate) were inadvertently introduced during the post-data collection updating process. The detailed range specifications are provided

in a later section; items for which values outside the hard range were accepted (with respondent confirmation) are also noted.

Logic Edits

Logic edits were used to examine the relationships between responses to ensure that they did not conflict with one another, or that the response to one item did not make the response to another item unlikely. If a discrepancy was discovered, the interviewer attempted to reconcile the difference while on the telephone with the respondent. These edits were performed on-line during the administration of the CATI application, and were readministered to ensure data integrity after manual data editing and data retrieval. The logic edit specifications, detailed in a later section, include the definition of the logic edits that were performed. Many of these edits were performed during CATI data collection. All of them were performed as a post-data collection effort.

Structural Edits

Because the relationships of database records were often dependent on values of variables contained in other database records, structural edit specifications were developed to ensure the structural integrity of the database (i.e., all variables that should exist do exist, and those that should not exist do not exist). For example, if there was a completed School Readiness interview for a preschooler, the data records that contain the developmental characteristics items must exist in the database. The edits for structural completeness are outlined below. It may be helpful to refer to appendix C, the database design diagram, when reviewing these edits.

Frequency and Cross-Tabulation Review

After data collection was completed, the frequencies of responses to all data items (both individually and in conjunction with related data items) were reviewed to ensure that appropriate skip patterns were followed. Members of the data preparation team checked each item to make sure the correct number of responses was represented. If a discrepancy was discovered, the problem case was identified and reviewed. If necessary, the audit trail for the interview, which provides a keystroke-by-keystroke record of the interview, was retrieved to determine the appropriate response. If the audit trail revealed no additional information, either a data clarification callback was made or the item was coded as "not ascertained."

Frequency Review of Text Items

The "Other, Specify" open-ended text responses were reviewed to determine if they should have been coded into one of the existing code categories. If so, the recoding was completed. If there was a response that occurred a substantial number of times for a particular text item, a new code was created, and the responses were recoded. In the NHES:93, additional response categories were added to some items in the School Safety and Discipline interview. Specific items are the main reason the child attends his/her current school (P13); when, if ever, it would be all right with parents for the child to smoke cigarettes (PY93); and when, if ever, it would be all right with parents for the child to drink alcohol (PY95).

Additional categories that were not part of the original questions are shown in italics in the questionnaires contained in the data file user's manuals.

Problem Areas and Suggestions for Improvements in Future Surveys

In this section, we discuss some of the problems identified during data editing in the NHES:93. Typically, the items described below were identified through CATI comments by interviewers, problem sheets completed by interviewers, or post-CATI editing or frequency review. All of these problems required updates to the data base, and some required that an interview be "cleaned out" and refielded.

1. Inaccuracies in the enumeration of household members (S6) occasionally caused problems in the administration of the parent/guardian questions. These inaccuracies included omission of household members, inclusion of nonhousehold members, and erroneous information about the household members listed. These problems were corrected in data preparation through the review of interview comments in the CATI system and the processing of problem sheets. Confirming all information on the enumeration screen would improve data quality and interview administration.

2. Similarly, there were problems with the relationship questions in the School Readiness interview (R10-R13), which were identified in post-interview editing. A summary screen after this series would assist the interviewer and the respondent in recording the correct household composition. Again, this would improve the administration of the parent/guardian questions.

3. After collecting the date of birth (R1/P1), a question should be added to future surveys asking the respondent to confirm the age of the child as calculated using the date of birth. If the respondent does not confirm the calculated age, the date of birth question should be reasked. This would prevent the need for updates to the child's age in the data preparation process. This is particularly important because the child's age often drives skip patterns.

4. Under the current design, the interviewer can only change parent respondents at the date of birth question, the enrollment question, or the grade question. Data preparation staff had to "clean out" interviews when the respondent wanted to have the other parent complete the interview after these "change points" had passed. More emphasis during interviewer training on changing respondents during extended interviews is necessary. Perhaps an interrupt module could be implemented to allow automated respondent switching at certain points in the interview.

5. There was some confusion about the definition of school enrollment at questions S6 and R4/P4 ("Is {child} attending or enrolled in school?"), especially in the School Readiness population. Respondents sometimes reported that their child was "enrolled" when their child was "enrolled" to attend kindergarten in the fall. This was identified in CATI comments, and required updating information in data preparation. This issue was included in question-by-question specifications for subsequent NHES surveys. Interviewers could back up to correct the enrollment item if the respondent indicated at the grade question that the child was not currently in a grade, but enrolled for the fall.

6. In the School Readiness interview, when the Screener respondent and the extended respondent were the same, the enrollment information collected in the Screener was copied into the extended interview. The home school item (R5/P5) is one of these copied questions. This question is only

asked of 5-, 6-, and 7-year-olds. However, if there was a discrepancy between the screener age and the R1/P1 calculated age, S4 may or may not have been asked when R5/P5 should have or should not have been asked. Program adjustments would be helpful in dealing with age-driven skips for such items and would prevent the need for data preparation updates.

7. The number of edit failures was greater for the question on whether there was an educational program in an early childhood care program (R40/R44) than in most other items (greater than one percent edit failure). This item was dropped from subsequent NHES collections about early childhood programs.

8. At R137/P99 and R149/P107 (highest grade of school attended by mother/father), there was confusion about how to code "did not go to school at all." This response was coded as 1 during post-data collection editing. However, training on this issue or a note on the screen to the interviewer would avoid unnecessary editing.

9. At R141/P103 and R153/P111 (number of hours mother/father work per week), a number of responses exceeded the upper range limit of 80. An extension to 99 would accommodate most of these exceptions.

10. Additional training on the definition of emancipated youth would be helpful, if future NHES instruments have specific paths for such cases. It was found in data editing that there were some emancipated youths who had "parent age" adults living in the household with them.

11. There was some confusion associated with when to ask certain items in the parent/guardian section, especially when there were no parents/guardians living in the household. This problem was identified in post-interview editing. For NHES:95, these issues were clarified by amending the CATI indicators of parents or guardians in the household (HHMOM and HHDAD).

Logic and Range Edit Specifications

The logic edits and range are defined below for each component. After the specification of each of the edits used during CATI there is a review of the number of times the edit was invoked. The number eligible for each edit represents the population of people who had nonmissing data for all variables involved in the edit. Please note that counts of occurrences are available only for the logic edits done during CATI, not for range edits and post-data collection edits.

Only four of the online edits resulted in greater than one percent failure. They are S10 and S10/R1 (grade equivalent vs. calculated age), R40/R44 (educational programs in care programs), and R55 (when child is expected to start kindergarten vs. calculated age). Two edits, R7/P7 (grade equivalent) vs. calculated age and P15/P16 for emancipated youth (highest grade >= lowest grade), were never triggered.

Screener Items

S5a. *Anyone 18 or younger in household*

If S3 = 3, then S5a = 1

*** Of the 812 households eligible, 1 triggered this edit (0.12%).

S9. *Grade or year child is attending vs. Screener AGE vs. calculated age* (DOBYY (R1/P1) and interview date)

Age < 5, grade -1, N, T, K, P, 13, 14
Age = 5, grade -1, N, T, K, P, 1, 13, 14
Age = 6, grade -1, N, T, K, P, 1, 2, 13, 14
Age = 7, grade -1, T, K, P, 1, 2, 3, 13, 14
Age = 8, grade -1, 1, 2, 3, 4, 13, 14
Age = 9, grade -1, 2, 3, 4, 5, 13, 14
Age = 10, grade -1, 3, 4, 5, 6, 13, 14
Age = 11, grade -1, 4, 5, 6, 7, 13, 14
Age = 12, grade -1, 5, 6, 7, 8, 13, 14
Age = 13, grade -1, 6, 7, 8, 9, 13, 14
Age = 14, grade -1, 7, 8, 9, 10, 13, 14
Age = 15, grade -1, 8, 9, 10, 11, 13, 14
Age = 16, grade -1, 9, 10, 11, 12, 13, 14
Age = 17, grade -1, 10, 11, 12, 13, 14, 15, 16
Age = 18, grade -1, 11, 12, 13, 14, 15, 16
Age = 19, grade -1, 11, 12, 13, 14, 15, 16
Age = 20, grade -1, 11, 12, 13, 14, 15, 16

*** Of the 30,415 people eligible for the Screener age edit, 244 triggered it (0.80%).

*** Of the 21,441 people eligible for the calculated age edit, 152 triggered it (0.71%).

S10. *Grade equivalent*
vs. Screener AGE
vs. calculated age
(DOBYY (R1/P1) and interview date)
Age < 5, grade -1, N, T, K, P, 13
Age = 5, grade -1, N, T, K, P, 1, 13
Age = 6, grade -1, N, T, K, P, 1, 2, 13
Age = 7, grade -1, T, K, P, 1, 2, 3, 13
Age = 8, grade -1, 1, 2, 3, 4, 13
Age = 9, grade -1, 2, 3, 4, 5, 13
Age = 10, grade -1, 3, 4, 5, 6, 13
Age = 11, grade -1, 4, 5, 6, 7, 13
Age = 12, grade -1, 5, 6, 7, 8, 13
Age = 13, grade -1, 6, 7, 8, 9, 13
Age = 14, grade -1, 7, 8, 9, 10, 13
Age = 15, grade -1, 8, 9, 10, 11, 13
Age = 16, grade -1, 9, 10, 11, 12, 13
Age = 17, grade -1, 10, 11, 12, 13
Age = 18, grade -1, 11, 12, 13
Age = 19, grade -1, 11, 12, 13
Age = 20, grade -1, 11, 12, 13

*** Of the 264 people eligible for the Screener age edit, 24 triggered it (9.09%).
*** Of the 210 people eligible for the calculated age edit, 6 triggered it (2.86%).

S12. *Relationship between child and most knowledgeable parent/guardian*
If S12 = 1, 2, then parent/guardian age > child's age + 15
If S12 = 4, then parent/guardian age > child's age + 35
If S12 = 1, then gender = F
If S12 = 2, then gender = M
Only 1 household member can have S12 = 1
Only 1 household member can have S12 = 2

*** Of the 23,189 people eligible for the AGE part of this edit, 58 triggered it (0.25%).
*** Of the 24,733 people eligible for the sex part of this edit, 174 triggered it (0.70%).

Items Common to Both School Readiness and School Safety and Discipline

R1/P1. *Month and year of child's birth*
Month: 1 - 12 (hard range)
Year: School Readiness: 1983 - 1989
School Safety and Discipline: 1972 - 1984
(value outside of range goes to CLOSE1; all entries require confirmation)
AGE92 (calculated age) = 92 - R1/P1 year

R6/P6. *Grade or year child is attending*

Age < 5, grade -1, N, T, K, P, 13, 14
Age = 5, grade -1, N, T, K, P, 1, 13, 14
Age = 6, grade -1, N, T, K, P, 1, 2, 13, 14
Age = 7, grade -1, T, K, P, 1, 2, 3, 13, 14
Age = 8, grade -1, 1, 2, 3, 4, 13, 14
Age = 9, grade -1, 2, 3, 4, 5, 13, 14
Age = 10, grade -1, 3, 4, 5, 6, 13, 14
Age = 11, grade -1, 4, 5, 6, 7, 13, 14
Age = 12, grade -1, 5, 6, 7, 8, 13, 14
Age = 13, grade -1, 6, 7, 8, 9, 13, 14
Age = 14, grade -1, 7, 8, 9, 10, 13, 14
Age = 15, grade -1, 8, 9, 10, 11, 13, 14
Age = 16, grade -1, 9, 10, 11, 12, 13, 14, 15, 16
Age = 17, grade -1, 10, 11, 12, 13, 14, 15, 16
Age = 18, grade -1, 11, 12, 13, 14, 15, 16
Age = 19, grade -1, 11, 12, 13, 14, 15, 16
Age = 20, grade -1, 11, 12, 13, 14, 15, 16

*** Of the 21,432 people eligible for this edit, 24 triggered it (0.11%).

R7/P7. *Grade equivalent*

Age < 5, grade -1, N, T, K, P, 13
Age = 5, grade -1, N, T, K, P, 1, 13
Age = 6, grade -1, N, T, K, P, 1, 2, 13
Age = 7, grade -1, T, K, P, 1, 2, 3, 13
Age = 8, grade -1, 1, 2, 3, 4, 13
Age = 9, grade -1, 2, 3, 4, 5, 13
Age = 10, grade -1, 3, 4, 5, 6, 13
Age = 11, grade -1, 4, 5, 6, 7, 13
Age = 12, grade -1, 5, 6, 7, 8, 13
Age = 13, grade -1, 6, 7, 8, 9, 13
Age = 14, grade -1, 7, 8, 9, 10, 13
Age = 15, grade -1, 8, 9, 10, 11, 13
Age = 16, grade -1, 9, 10, 11, 12, 13
Age = 17, grade -1, 10, 11, 12, 13
Age = 18, grade -1, 11, 12, 13
Age = 19, grade -1, 11, 12, 13
Age = 20, grade -1, 11, 12, 13

*** Of the 175 people eligible for this edit, none triggered it.

R141/P103. *Hours per week mom usually works for pay*

1 - 80 (hard range)
1 - 50 (soft range)

R153/P111. *Hours per week dad usually works for pay*
(values outside hard range accepted if confirmed by respondent; actual range 1-99)[13]
1 - 80 (hard range)
1 - 50 (soft range)

R161/P117. *Number of bedrooms in the house*
0 - 10 (hard range)
1 - 6 (soft range)
Note: Data collection range was 0 - 6; post-CATI edit was 0 - 10.

R164/P119. *Number of additional telephones for home use*
0 - 9 (hard range)
0 - 3 (soft range)

R166/P121. *Amount of time telephone was out of service in past 12 months*
Days: 1 - 365 (hard range)
1 - 120 (soft range)
Weeks: 1 - 52 (hard range)
1 - 16 (soft range)
Months: 1 - 12 (hard range)
1 - 4 (soft range)
All responses to this question were converted to DAYS using 28 days in a month and 7 days in a week.

R167/P122. *Zip Code*
The telephone exchange of the sampled case was matched to the Donnelley tape to identify the first three digits of the Zip code. About 16 percent of telephone exchanges did not have a 3-digit ZIP value on the Donnelley tape. The 3-digit Zip was used as an edit check when respondents reported their Zip codes. Edit allowed respondent to verify their response if it did not match the Donnelly 3-digit Zip code.

School Readiness Items

R10. *How person is related to child*
If R10 = 1, 2, then person's age > subject's age + 15
If R10 = 4, then person's age > subject's age + 35
If R10 = 1, then gender = F
If R10 = 2, then gender = M
Only 1 household member with S12/R10 = 1
Only 1 household member with S12/R10 = 2
*** Of the 26,539 people eligible for the age edit, 84 triggered it (0.32%).
*** Of the 26,535 people eligible for the mother/father sex edit, 68 triggered it (0.26%).
*** Of the 26,535 people eligible for the multiple mother/father edit, 191 triggered it (0.72%).

[13] A notation appears if the actual range in the data set is greater than the edit range. Out-of-range values were confirmed by respondents. If there is no notation, all values are within the edit range.

R34. *Age of child when first attended Head Start program*

Years: 2 - current age [maximum 6 years] (hard range)
3 - 5 years (soft range)
Months: 0 - 11
R34 age cannot exceed age as of interview date

R35. *How long child attends/attended Head Start*

1 - 5 (hard range)
If age as of interview date < 5, R35 must be < 4
If age as of interview date < 3, R35 must be < 2

R38. *Age of child when first attended any nursery school, prekindergarten, preschool, or a day care center*

Years: 0 - 7 [maximum current age]
Months: 0 - 11 [maximum current age]
If years = 0, then months > 0
R38 age cannot exceed age as of interview date

R39. *How long child attended any nursery school, prekindergarten, preschool, or day care center*

1 - 5 (hard range)
Must be <= age as of interview date + 1

Edit check verified that response was <= current age + 1 since response categories do not map directly to ages. For example, if child is 2 years old and has spent entire life in daycare center, response of category 3 (2 years, less than 2) would be acceptable.

R41. *Number of Head Start programs, nursery schools, prekindergartens, preschools, or day care centers child goes to now*

1 - 4 (hard range)
1 - 2 (soft range)

R44. *Educational program in current setting*

If R44 = 1, then R40 = 1
*** Of the 2,402 people eligible for this edit, 124 triggered it (5.16%).

R45. *Number of days each week child goes to the program*

1 - 7 (hard range)

R46. *Number of hours each week child goes to the program*

1 - 70 (hard range)
1 - 50 (soft range)
R46/R45 cannot exceed 9 hours per day (could be overridden with confirmation)

R47. *Does child attend a full- or part-day program*[14]

If R47 NE 1, R46/R45 <= 5 (if child is not in a full-day program, average hours per day must be < = 5 hours).

If R47 NE 2, R46/R45 >= 5 (if child is not in a part-day program, average hours per day must be >= 5 hours).

R48. *Number of children in room or group at program*

1 - 40 (hard range)

5 - 20 (soft range)

R49. *Number of adults in room or group at program*

(values outside the hard range confirmed with respondent; actual range 1-22)

1 - 10 (hard range)

1 - 4 (soft range)

R55. *When child is expected to start kindergarten*

If age = 3, R55 = 1, 4 - 9

If age = 4, R55 = 1 - 9

If age = 5, R55 = 1 - 9

If age = 6, R55 = 1 - 3

If age = 7, R55 = 1 - 3

*** Of the 4,425 people eligible for this edit, 377 triggered it (8.52%).

R59. *Age of child when first started kindergarten*

Years: 3 - 7 (hard range)

4 - 6 (soft range)

Months: 0 - 11

R64. *Number of hours each week spent in kindergarten*

(values outside hard range accepted if confirmed by respondent; actual range 5-50)

6 - 45 (hard range)

10 - 30 (soft range)

R76. *Number of times changed schools from the start of kindergarten/first grade until now*

0 - 10 (hard range)

0 - 3 (soft range)

R80A. *Has child skipped grade*

If grade = 2 and age => 7, then R80A = 2

*** Of the 2,151 people eligible for this edit, 14 triggered it (0.65%).

R80B. *Grades skipped*

Cannot equal or exceed current grade

---

[12]See the data anomalies chapter of the *National Household Education Survey of 1993: School Readiness Users' Manual* (Brick et al. 1994) for further discussion of items R44-R47.

R82. *Grade or grades repeated*

Cannot be greater than current grade. (Display only response categories up to current grade.)

R88. *Age of child when began reading simple, whole sentences*

Years: 3 - 8 [maximum current age] (hard range)
4 - 6 [maximum current age] (soft range)
Months: 0 - 11

R92. *Number of hours of television or video tape watching during the weekdays.*

R92a. *Before 8 am*

(values outside hard range accepted if confirmed by respondent; actual range 0-5)
Hours: 0 - 3 (hard range)
Minutes: 0 - 59 (hard range)
If R92a = 95, then R92b-d and R93 = -1 (skip pattern check)

R92b. *Between 8 am and 3 pm*

If grade = 1, 2, 3, then R92b <= 2 hours (soft range)
If full-day K, then R92b <= 2 hours (soft range)
Else, 0 - 7 hours (hard range)
Minutes: 0 - 59 (hard range)

R92c. *Between 3 pm and dinner time*

(values outside hard range accepted if confirmed by respondent; actual range 0-7)
hours: 0 - 4 (hard range)
minutes: 0 - 59 (hard range)

R92d. *After dinner time*

(values outside hard range accepted if confirmed by respondent; actual range 0-10)
hours: 0 - 5 (hard range)
minutes: 0 - 59 (hard range)

R93. *Number of hours of television or video tape watching during the weekends.*

R93a. *Saturdays*

hours: 0 - 16 (hard range)
hours: 0 - 6 (soft range)
minutes: 0 - 59 (hard range)

R93b. *Sundays*

hours: 0 - 16 (hard range)
hours: 0 - 6 (soft range)
minutes: 0 - 59 (hard range)

R113. *Number of days child ate breakfast*

0 - 7 (hard range)

R114. *Number of days someone fixed a hot meal for the child*

0 - 7 (hard range)

R115. *Number of times family ate dinner together*

0 - 7 (hard range)

R120. *How long child lived apart from mother*

years: 0 - 5 [maximum current age] (hard range)
months: 0 - 11 (hard range)
If years = 0 then months > 0
If years = 5 then months = 0
If years => 1 then months = 0 - 11

NOTE: On-line edit allowed entry of up to 24 months if the years = 0. Entries of > 12 months were converted to years and months post-interview. This note applies to R120 through R132 and to R146.

R124. *Mother worked full-time outside of the house for pay*

0 - 5 years [maximum current age] (hard range)
months: 0 - 11 (hard range)
If years = 0 then months > 0.
If years = 5 then months = 0
If years => 1 then months = 0 - 11

R126. *Family had serious financial problems*

0 - 5 years [maximum current age] (hard range)
months: 0 - 11 (hard range)
If years = 0 then months > 0
If years = 5 then months = 0
If years => 1 then months = 0 - 11

R128. *Family received food stamps*

0 - 5 years [maximum current age] (hard range)
months: 0 - 11 (hard range)
If years = 0 then months > 0
If years = 5 then months = 0
If years => 1 then months = 0 - 11

R130. *Family received welfare or AFDC*

0 - 5 years [maximum current age] (hard range)
months: 0 - 11 (hard range)
If years = 0 then months > 0
If years = 5 then months = 0
If years => 1 then months = 0 - 11

R131. *Number of times family moved*

0 - 15 (hard range)
0 - 5 (soft range)

R132. *Age of child when mother first lived with him/her*

Years: 0 - current age (hard range)
Months: 0 - 11 (hard range)
If years = 0 then months > 0
If years => 1 then months = 0 - 11

R142. *Months mother has worked in past year*

0 - 12 (hard range)

R146. *Age of child when father first lived with him/her*

Years: 0 - 9 [maximum current age] (hard range)
Months: 0 - 11 (hard range)
If years = 0 then months > 0
If years >0 then months 0 - 11

School Safety and Discipline Items

S13. *Type of mother/father (School Safety and Discipline Parent only, RESPTYPE = P1, P2)*

If S12 = 1 & S13 = 1, then P8 = 1
If S12 = 1 & S13 = 2, then P8 = 1
If S12 = 1 & S13 = 3, then P8 = 2
If S12 = 1 & S13 = 4, then P8 = 3
If S12 = 2 & S13 = 1, then P9 = 1
If S12 = 2 & S13 = 2, then P9 = 1
If S12 = 2 & S13 = 3, then P9 = 2
If S12 = 2 & S13 = 4, then P9 = 3

P15. *Lowest grade at child's school*

P15 ♋ P6/P7

P16. *Highest grade at child's school*

P16 ♍ P6/P7
P16 ♍ P15

*** Of the 10,130 P1 interviews eligible for P15/P16 edit, 53 triggered the edit (0.52%).

*** Of the 2,572 P2 interviews eligible for P15/P16 edit, 7 triggered the edit (0.27%).

*** No emancipated youths triggered this edit.

P18. *Number of students at child's school*

1 - 5, 10 - 2000 (hard range)
1 - 5, 10 - 599 (soft range)

P20. *Gender composition of school*

If gender = M, then P20 = 1, 2
If gender = F, then P20 = 1, 3

Y52. *Items youth brought to school*

If Y51 = 1 then Y52 a, b, c, d, e, f, g, h, and i cannot all = 2

P70. *Number of days child missed school*

(values outside the hard range accepted if confirmed by respondent; actual range 0-24)
0 - 20 (hard range)
0 - 8 (soft range)

P74. *Child ever expelled from school*

If P13 = 3 then P74 = 1

*** Of the 1,948 people eligible for this edit, none triggered it.

P81. *Number of times child has moved*

values outside hard range accepted if confirmed by respondent; actual range 0-15)
0 - 10 (hard range)
0 - 5 (soft range)

PY93. *Time child can smoke*

1 - 8, current age - 35, 91 (hard range)
1 - 8, current age - 21, 91 (soft range)

PY95. *Time child can drink alcohol*

1 - 8, current age - 35, 91 (hard range)
1 - 8, current age - 21, 91 (soft range)

Edits for Structural Completeness

In order to understand the structural editing process in the NHES:93, it is necessary to understand the nature of the CATI data base. The CATI system used for the NHES:93 employed a hierarchical data base structure containing records for administrative purposes and records that held responses to interviews (see appendix C for a graphical presentation). The BASE segment was the key segment for each sampled telephone number. CLUS contained information on the screening of the primary numbers of the clusters (100-banks). Four administrative segments contained records associated with each interview: SKED, CATI scheduler records; SURV, records for each contact attempt; NIRF, noninterview (refusal, language, or problem) records; and MESS, messages left by one interviewer for another. Screener records were SCRN, containing household-level responses to Screener items; and ENUM, containing person-level information collected in the Screener. BASM was the adminstrative record for each extended interview. School Readiness items were contained on REDY, which contained questionnaire items asked about all children; PRES, which contained items asked only about preschoolers, PRIM, which contained items asked only about primary school children or kindergartners; SCHL, which contained the school-related information; RELN, which contained information on relationships within the household, DEMO, which contained demographic characteristics, MAMA, which contained information on the child's mother, PAPA, which contained information on the child's father; and HOME, which contained household characteristics. For School Safety and Discipline parent interviews, data were stored on PAR1 for parents of 6th through 12th graders and on PAR2 for parents of 3rd to 5th graders. DEMO, MAMA, PAPA, and HOME (described above)

were also used for SS&D parent interviews. The YUTH segment contained responses to the SS&D youth interview, and EMAN contained responses from emancipated youth.

The structural edits defined below were run against completed extended interviews only (BASM.MAINRSLT = C1,C2,CY,CE,CN,CK,CS,CH). The edits can be grouped into three logical categories: edits that verify interview completeness, edits that confirm the presence of appropriate person records, and edits that verify parent relationships in the household.

Interview Completeness

These edits confirm the completeness of the database. In other words, if there is a completed interview, all of the appropriate records associated with that type of interview exist.

If SCRN.SCRNRSLT = CD, CR, CB
then there must be a BASM record with SELECTEX = SR, SP, SY, or SE.

Each completed interview (MAINRSLT = C1,C2,CE,CY,CH,CN,CK,CS) with BASE.PHONTYPE = 1 (prime number) must have one and only one BASE record, CLUS record, and SCRN record.

Each completed interview with BASE.PHONTYPE = 2 (secondary number) must have one and only one BASE record and SCRN record.

Each completed interview with BASM.MAINRSLT = C1,C2,CY,CH,CN,CK,CS,CE must have one and only one HOME record.

If BASM.MAINRSLT = C1
then there must be a DEMO record & a PAR1 record.

If BASM.MAINRSLT = C2
then there must be a DEMO record & a PAR2 record.

If BASM.MAINRSLT = CE
then there must be a DEMO record with DEMO.MOMHOME = 4 and DEMO.DADHOME = 4, an EMAN record, and a YUTH record.

If BASM.MAINRSLT = CH
then there must be a DEMO record and a REDY record.

If BASM.MAINRSLT = CK
then there must be a DEMO record, a SCHL record, and a REDY record.

If BASM.MAINRSLT = CN
then there must be a DEMO record, a PRES record, and a REDY record.

If BASM.MAINRSLT = CS
then there must be a DEMO record, a SCHL record, a PRIM record, and a REDY record.

If BASM.MAINRSLT = CY

then there must be a DEMO record, a YUTH record, a PAR1 record, and another BASM record for the youth with BASM.MAINRSLT = C1.

If BASM.MAINRSLT = C1
then there must be NO PAR2, EMAN, REDY, PRES, SCHL, or PRIM record.

If BASM.MAINRSLT = C2
then there must be NO PAR1, YUTH, EMAN, REDY, PRES, SCHL, or PRIM record.

If BASM.MAINRSLT = CY
then there must be NO PAR2, REDY, PRES, SCHL, or PRIM record.

If BASM.MAINRSLT = CE
then there must be NO PAR1, PAR2, REDY, PRES, SCHL, or PRIM record.

If BASM.MAINRSLT = CN
then there must be NO PAR1, PAR2, YUTH, EMAN, SCHL, or PRIM record.

If BASM.MAINRSLT = CK
then there must be NO PAR1, PAR2, YUTH, EMAN, PRES, or PRIM record.

If BASM.MAINRSLT = CS
then there must be NO PAR1, PAR2, YUTH, EMAN, or PRES, record.

If BASM.MAINRSLT = CH
then there must be NO PAR1, PAR2, YUTH, EMAN, PRES, SCHL, or PRIM record.

Appropriate Person Records

Every completed interview must have the appropriate associated person records. This includes person records for the subject and the respondent, as well as the mother, the father, and any other household members.

If BASM.MAINRSLT = CH, CN, CK, CS
then there must be a RELN record with RELNID = child's ENUMID || ENUM.PERSNUM of every other ENUM except child's ENUM.PARNNUM.

Every BASM must have an ENUM record with ENUM.PERSNUM = BASM.ENUMNUM and ENUM.ELIGFLG = 1.

If ENUM.PARNNUM ^= -1
then there must be an ENUM with ENUM.PERSNUM = ENUM.PARNNUM of child's ENUM and ENUM.SEX = ENUM.PARSEX of child's ENUM.

If DEMO.MOMNUM ^= -1
then there must be an ENUM record with ENUMID = first 8 digits of DEMOID || DEMO.MOMNUM.

If DEMO.DADNUM ^= -1
then there must be an ENUM record with ENUMID = first 8 digits of DEMOID || DEMO.DADNUM.

If there is a RELN
then there must be an ENUM record with ENUMID = first 10 digits of RELNID and
another ENUM where ENUMID = first 8 digits of RELNID || 11th and 12th digits of RELNID.

If SCRN.SCRNRSLT ^= -1
then there must be an ENUM where ENUM.PERSNUM = SCRESPX and ENUM.SCRESP = X.

Parent Relationships

Finally, parent relationships can be checked using the following edits.

If (any RELN.RELATION = 1 & RELN.FEMGUARD = 1 or 2) or (ENUM.PARNRELN = 1 & ENUM.PARNTYPE = 1 or 2)
then DEMO.REALMOM = 1
else DEMO.REALMOM = -1.

If (any RELN.RELATION = 2 & RELN.MALGUARD = 1 or 2) or (ENUM.PARNRELN = 2 & ENUM.PARNTYPE = 1 or 2)
then DEMO.REALDAD = 1
else DEMO.REALDAD = -1.

If any RELN.RELATION = 1 or ENUM.PARNRELN = 1 or DEMO.MOMHOME = 1, 2, or 3
then there must be an ENUM where ENUM.PERSNUM = DEMO.MOMNUM.

If any RELN.RELATION = 2 or ENUM.PARNRELN = 2 or DEMO.DADHOME = 1, 2, or 3
then there must be an ENUM where ENUM.PERSNUM = DEMO.DADNUM.

If DEMO.HHDAD = 1
then there must be a PAPA record where PAPAID = BASEID||DADNUM.

If DEMO.HHDAD ^= 1 & DEMO.HHMOM ^= 1 & ENUM.PARSEX = M & BASM.MAINRSLT ^= CE
then there must be a PAPA record where PAPAID = BASEID||PARNNUM.

If there is a PAPA for an ID
then there must be a DEMO with DEMO.DADNUM = the last 2 digits of PAPAID or there must be an ENUM record with the last 2 digits of ENUMID = ENUM.PARNNUM of the child's ENUM record.

If DEMO.HHMOM = 1
then there must be a MAMA record with MAMAID=BASEID||MOMNUM.

If DEMO.HHMOM ^= 1 & DEMO.HHDAD ^= 1 & ENUM.PARSEX = F & BASM.MAINRSLT ^= CE
then there must be a MAMA record where MAMAID=BASEID||PARNNUM.
If there is a MAMA record for an ID
then there must be a DEMO record with DEMO.MOMNUM = the last 2 digits of MAMAID or there must be an ENUM record with the last 2 digits of the ENUMID = ENUM.PARNNUM of the child's ENUM record.

If any BASM.MAINRSLT = CN, CK, CS, CH and its associated DEMO.HHMOM = 1 or (DEMO.HHMOM = -1 & DEMO.HHDAD = -1 & ENUM.PARSEX = F) in a household
then there must be a MAMA record with MOMSTAT ^= -1 & MOMLANG ^= -1 & ((MOMLANG ^= 1 & MOMSPEAK ^= -1) | (MOMLANG = 1 & MOMSPEAK = -1)) & MOMMTHS ^= -1
else there must be a MAMA record with MOMSTAT = -1 & MOMLANG = -1 & MOMSPEAK = -1 & MOMMTHS = -1.

If any BASM.MAINRSLT = CN, CK, CS, CH and its associated DEMO.HHDAD = 1 or (DEMO.HHMOM = -1 & DEMO.HHDAD = -1 & ENUM.PARSEX = M) in a household
then there must be a PAPA record with DADLANG ^= -1 & ((DADLANG ^= 1 & DADSPEAK ^= -1) | (DADLANG = 1 & DADSPEAK = -1))
else there must be a PAPA record with DADLANG = -1 & DADSPEAK = -1.

If DEMO.BIRTHMOM = 1
then there must be a RELN with RELN.RELATION = 1 & RELN.FEMGUARD = 1 or
there must be an ENUM with ENUM.PARNRELN = 1 & ENUM.PARNTYPE = 1.

If (any RELN.RELATION = 1 & RELN.FEMGUARD = 1) or (ENUM.PARNRELN = 1 & ENUM.PARNTYPE = 1)
then DEMO.BIRTHMOM = 1.

If DEMO.BIRTHDAD = 1
then there must be a RELN with RELN.RELATION = 2 & RELN.MALGUARD = 1 or
there must be an ENUM with ENUM.PARNRELN = 2 & ENUM.PARNTYPE = 1.

If (any RELN.RELATION = 2 & RELN.MALGUARD = 1) or (ENUM.PARNRELN = 2 & ENUM.PARNTYPE = 1)
then DEMO.BIRTHDAD = 1.

IF DEMO.HHMOM= 1
then (ENUM.PARNRELN= 1) or (DEMO.MOMHOME = 1, 2, or 3) or (there must be a RELN with RELN.RELATION = 1).

If (any RELN.RELATION = 1) or ENUM.PARNRELN = 1 or DEMO.MOMHOME = 1, 2, or 3
then DEMO.HHMOM = 1.

If DEMO.HHDAD = 1
then (ENUM.PARNRELN= 2) or (DEMO.DADHOME = 1, 2, or 3) or (there must be a RELN with RELN.RELATION = 2).

If (any RELN.RELATION = 2) or ENUM.PARNRELN = 2 or DEMO.DADHOME = 1, 2, or 3 then DEMO.HHDAD = 1.

If ^missing(MOMNUM)
then there must be an ENUM record where ENUMID = first 8 digits of BASMID || DEMO.MOMNUM and ENUM.AGE <= child's AGE92 + 55 and ENUM.AGE >= child's AGE92 + 13.

If ^missing(DADNUM)
then there must be an ENUM record where ENUMID = first 8 digits of BASMID || DEMO.DADNUM and ENUM.AGE >= child's AGE92 + 13.

If ENUM.PARNRELN = 3-6
then there must be an ENUM where ENUMID = first 8 digits of BASMID || child's ENUM.PARNNUM and AGE >= 18.

If ENUM.PARNRELN = 4
then there must be an ENUM where ENUMID = first 8 digits of BASMID || child's ENUM.PARNNUM and AGE >= child's DEMO.AGE92 + 35.

For every BASM, there must be either (ENUM.PARNRELN = 1 and ENUM.PARNTYPE = 1) or at most one (RELN.RELATION = 1 and RELN.FEMGUARD = 1).

For every BASM, there must be either (ENUM.PARNRELN = 2 and ENUM.PARNTYPE = 1) or at most one (RELN.RELATION = 2 and RELN.MALGUARD = 1).

Skip Edits

Skip patterns are generally checked by reviewing frequencies for all data items. In addition to the standard structural and consistency edits, the following skip edits were programmed to identify skip errors for skips which were especially complex or dependent on numerous variables.

If (DEMO.ENROLL = 1 & DEMO.GRADE ^= 13,14) or
(DEMO.ENROLL = 2 & DEMO.AGE92 ^= 5,6,7) or
(DEMO.ENROLL = 2 & DEMO.AGE92 = 5,6,7 & DEMO.HOMESCHL = 2)
then DEMO.GRADEEQ = -1.

If MOMLOOK = 2 or missing or
(((MOMREAD = 1 or missing) or (MOMOTHER = 1 or missing)) and
MOMPUBL ^= 1 & MOMPRIV ^= 1 & MOMEMPL ^= 1 & MOMREL ^= 1 & MOMANSAD ^= 1)
then MOMACTY ^= -1
else MOMACTY = -1.

If DADLOOK = 2 or missing or
(((DADREAD = 1 or missing) or (DADOTHER = 1 or missing)) and

```
DADPUBL ^= 1 & DADPRIV ^= 1 & DADEMPL ^= 1 & DADREL ^= 1 & DADANSAD ^= 1)
        then DADACTY ^= -1
        else DADACTY = -1.

If BASM.MAINRSLT = CN, CK, CS, CH &
(HHDAD = 1 | (HHMOM ^= 1 & HHDAD ^= 1 & ENUM.PARSEX = M)) & BIRTHDAD ^= 1
        then DADKIDYR ^= -1.

If BASM.MAINRSLT = CN, CK, CS, CH &
(HHMOM = 1 | (HHMOM ^= 1 & HHDAD ^= 1 & ENUM.PARSEX = F)) & BIRTHMOM ^= 1
        then MOMKIDYR ^= -1.

If BASM.MAINRSLT = CN, CK, CS, CH &
(HHDAD ^= 1 & PARSEX = F) | BIRTHDAD = 1
        then DADKIDYR = -1.

If BASM.MAINRSLT = CN, CK, CS, CH &
(HHMOM ^= 1 & PARSEX = M) | BIRTHMOM = 1
        then MOMKIDYR = -1.

If BASM.MAINRSLT = CN, CK, CS, CH
        then HOME.TEFAMILY ^= -1.

If BASM.MAINRSLT = CN & KPSTART ^= 1
        then PRES.KPCOUNT ^= -1.

If BASM.MAINRSLT = CN, CK, CS, CH &
(REALMOM = -1 | (REALMOM = 1 & REALDAD = -1)) &
((PARSEX = F & HHDAD ^= 1) | PARSEX = M)
        then SEEPARN ^= -1.

If BASM.MAINRSLT = CN, CK, CS, CH
        then HLIVE ^= -1.

If BASM.MAINRSLT = CN, CK, CS, CH & HHMOM = 1
then MOMMARRY ^= -1.
```

References

Alexander, C.H., Sebold, J., and Pfaff, P. (1986). "Some results of an experiment with telephone sampling for the U.S.-National Crime Survey." *Proceedings of the American Statistical Association Section on Survey Research Methods*. Arlington, VA: American Statistical Association. p. 351-356.

Brick, J.M., and Waksberg, J. (1991). Avoiding sequential sampling with random digit dialing. *Survey Methodology*, 17(1):27-42.

Brick, J.M., Tubbs, E., Collins, M.A., Nolin, M.J., and Chandler, K. *Unit and Item Response, Weighting, and Imputaion Procedures in the 1993 National Household Education Survey*. NCES Working Paper 97-05. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Brick, J.M., Tubbs, E., Collins, M.A., Nolin, M.J., and Chandler, K. *Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey*. NCES Working Paper 97-02. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Nolin, M.J., Ha, P., Levinsohn, M., and Chandler, K. (1994a). *National Household Education Survey of 1993: School Readiness Data Files User's Manual*. NCES Publication No. 94-193. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Brick, J.M., Collins, M.A., Nolin, M.J., Ha, P., Levinsohn, M., and Chandler, K. (1994b). *National Household Education Survey of 1993: School Safety and Discipline Data Files User's Manual*. (NCES Publication No. 94-218.) Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Brick, J.M., Rizzo, L., Wernimont, J. (1996). *Reinterview Results for The School Readiness and School Safety and Discipline Components*. (NCES Publication No. 97-339.) Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Couper, M., Holland, L., and Groves, R. (1992). Developing Systematic Procedures for Monitoring in Centralized Telephone Facility. *Journal of Official Statistics*, 8(1), p. 63-75.

Groves, R., and Kahn, R. (1979). *Surveys By Telephone*. Academic Press, Inc.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Mohadjer, L. (1988). "Stratification of Prefix Areas for Sampling Rare Populations," Chapter 10 in *Telephone Survey Methodology*, edited by Groves, et al., New York.

Mohadjer, L., and West, J. (1992). *Effectiveness of Oversampling Blacks and Hispanics in the NHES Field Test*. (NCES Publication No. 92-104.) Washington, DC: U.S. Department of Education, National Center for Education Statistics.

National Commission on Children. (1991). *Speaking of Kids*. Washington, DC.

Nolin, M.J., and Chandler, K. (1996). *Use of Cognitive Laboratories and Recorded Interviews in the National Household Education Survey*. (NCES Publication No. 96-332.) Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Suchman, L., and Jordan, B. (1990). Interactional Troubles in Face-to-Face Interviews. *Journal of the American Statistical Association*, 85(409): 232-91.

Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of American Statistical Association*. 73(361): 40-46.

Waksberg, J. (1973). "The effect of stratification with differential sampling rates on attributes of subsets of the population." *Proceedings of the Social Statistics Section of the American Statistical Association*. Arlington, VA: American Statistical Association. p. 429-434.

APPENDIX A

MONITORING FORM

Monitoring Form

Interviewer _____ |_|_|_|_|_| |_|_| |_|_| |_|_|
NAME INITIALS MO DA YR

Monitor _____ |_|_|_|_|_|
NAME INITIALS

Project _____ |_|_|_|_|_|_|_|
NAME NUMBER

TIME
BEGIN |_|_| |_|_| AM
HR MIN
END |_|_| |_|_| AM
HR MIN

MONITORED IN THIS SESSION

Sample Characteristics:

	YES	NO
RDD	1	2
List HH	1	2
Business	1	2
CATI	1	2
Pretest	1	2
Other	1	2

(Specify)_____

Interview Type:

	YES	NO
Screener	1	2
Questionnaire	1	2
Data Retrieval	1	2
Refusal Conversion	1	2
Validation	1	2
Tracing	1	2
Prompt	1	2
Language	1	2
Other	1	2

(Specify)_____

Estimated Number of: Contacts |_|_| Questions Asked |_|_|_|

	N/A	NO PROB-LEM	NEEDS ATTENTION: Minor Difficulty	NEEDS ATTENTION: Major Difficulty	COMMENTS (IF DIFFICULTY NOTED, PROVIDE Q# AND COMMENT)	DISCUSSED 1=YES 2=NO
1. READING & GENERAL SKILLS a. Identifies self and reads intro clearly and without pausing.		1	2	3		1 2
b. Reads all appropriate phrases and answer categories		1	2	3		1 2
c. Follows skip and box instructions		1	2	3		1 2
d. Reads questions clearly with appropriate volume		1	2	3		1 2
e. Verifies spelling, address,						

BEST COPY AVAILABLE

phone numbers, and other data as needed		1	2	3		1	2
f. Adjusts pace of interview to accommodate respondent		1	2	3		1	2

Interviewer Signature

	N/A	NO PROBLEM	NEEDS ATTENTION: Minor Difficulty	NEEDS ATTENTION: Major Difficulty	COMMENTS (IF DIFFICULTY NOTED, PROVIDE Q# AND COMMENT)	DISCUSSED 1=YES 2=NO	
2. LISTENING SKILLS AND PROBING							
a. Listens to entire answer		1	2	3		1	2
b. Listens for what may not be said and probes		1	2	3		1	2
c. Probes unclear responses		1	2	3		1	2
d. Remains neutral when probing		1	2	3		1	2
3. RECORDING							
a. Records information accurately		1	2	3		1	2
b. Uses comments appropriately		1	2	3		1	2
c. Corrects coding errors		1	2	3		1	2
d. Uses control keys properly		1	2	3		1	2
e. Records result codes correctly		1	2	3		1	2
f. Moved through matrix and selection screens properly		1	2	3		1	2
g. Uses HH select screens properly and smoothly		1	2	3		1	2
4. HANDLING REFUSALS AND QUESTIONS							
a. Answers respondent questions and objections clearly, confidently, and briefly without hesitation		1	2	3		1	2
b. Offers verification number		1	2	3		1	2
5. TELEPHONE MANNER AND RELATIONSHIP WITH RESPONDENT							
a. Is pleasant, confident, and professional		1	2	3		1	2
b. Refrains from giving personal remarks or opinions		1	2	3		1	2

S1

BEST COPY AVAILABLE

c. Accepts emotions and sentiments without becoming personally involved		1	2	3		1 2

COMMENTS: _____

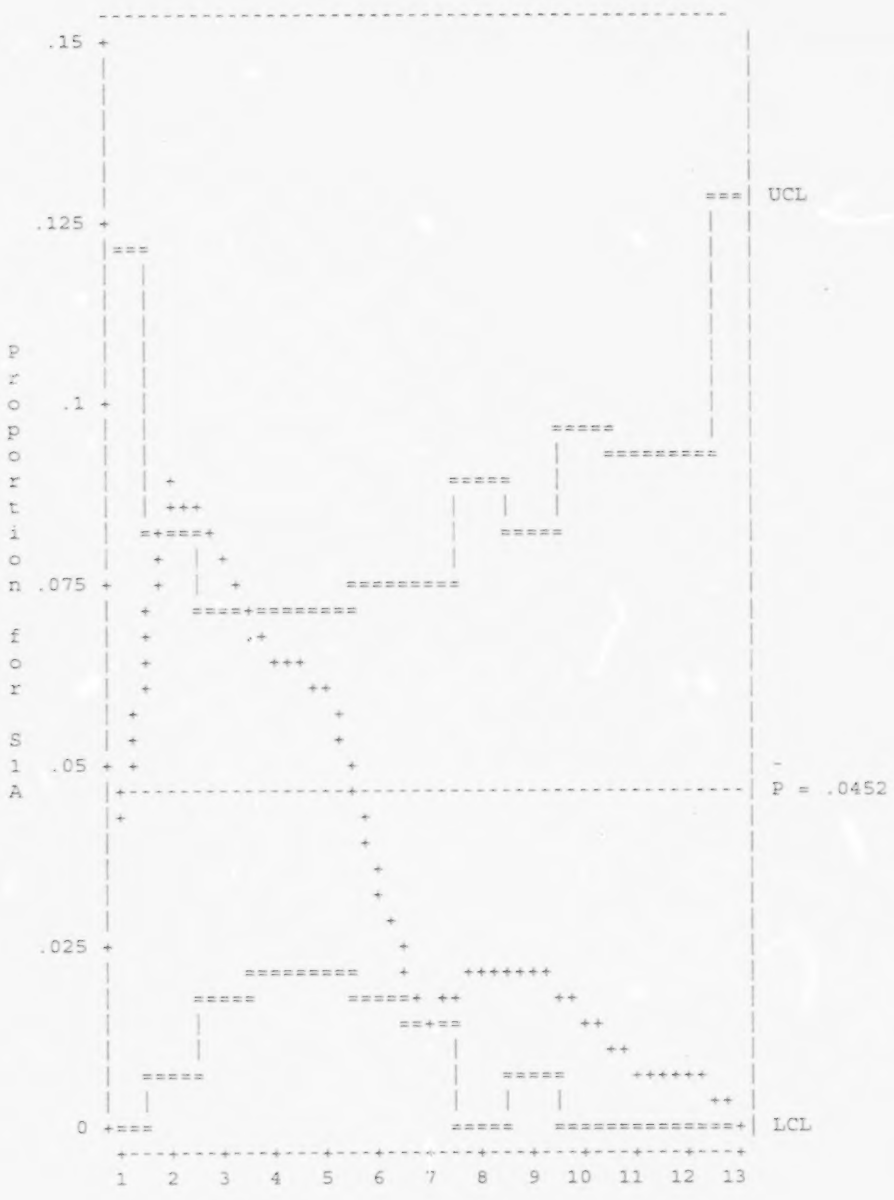
_____

_____

APPENDIX B

GRAPHS OF MONITORING RESULTS

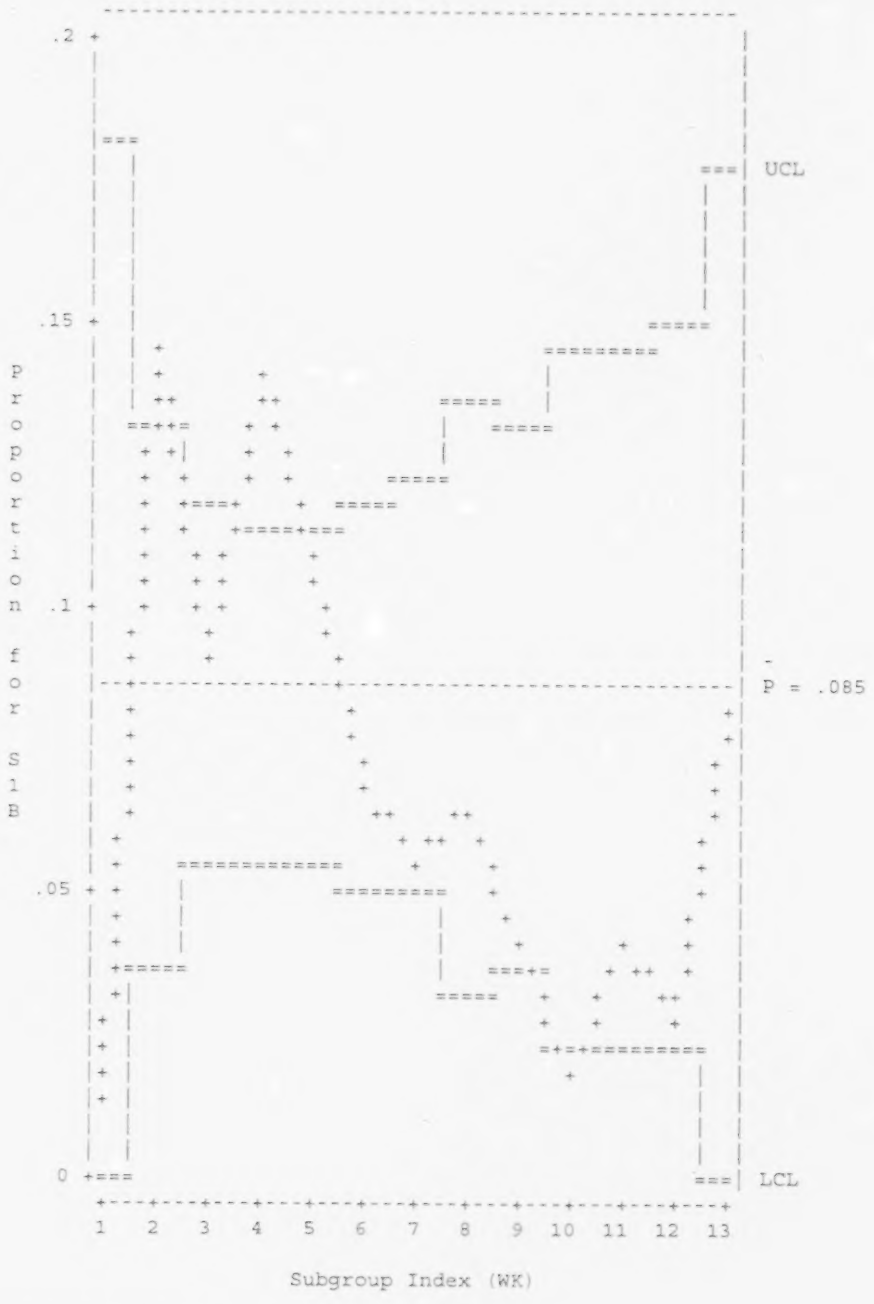
Figure 1.--Problems with item S1A, identifies self and reads introduction clearly and without pausing.



Subgroup Sizes: Min n=55 Max n=616

BEST COPY AVAILABLE

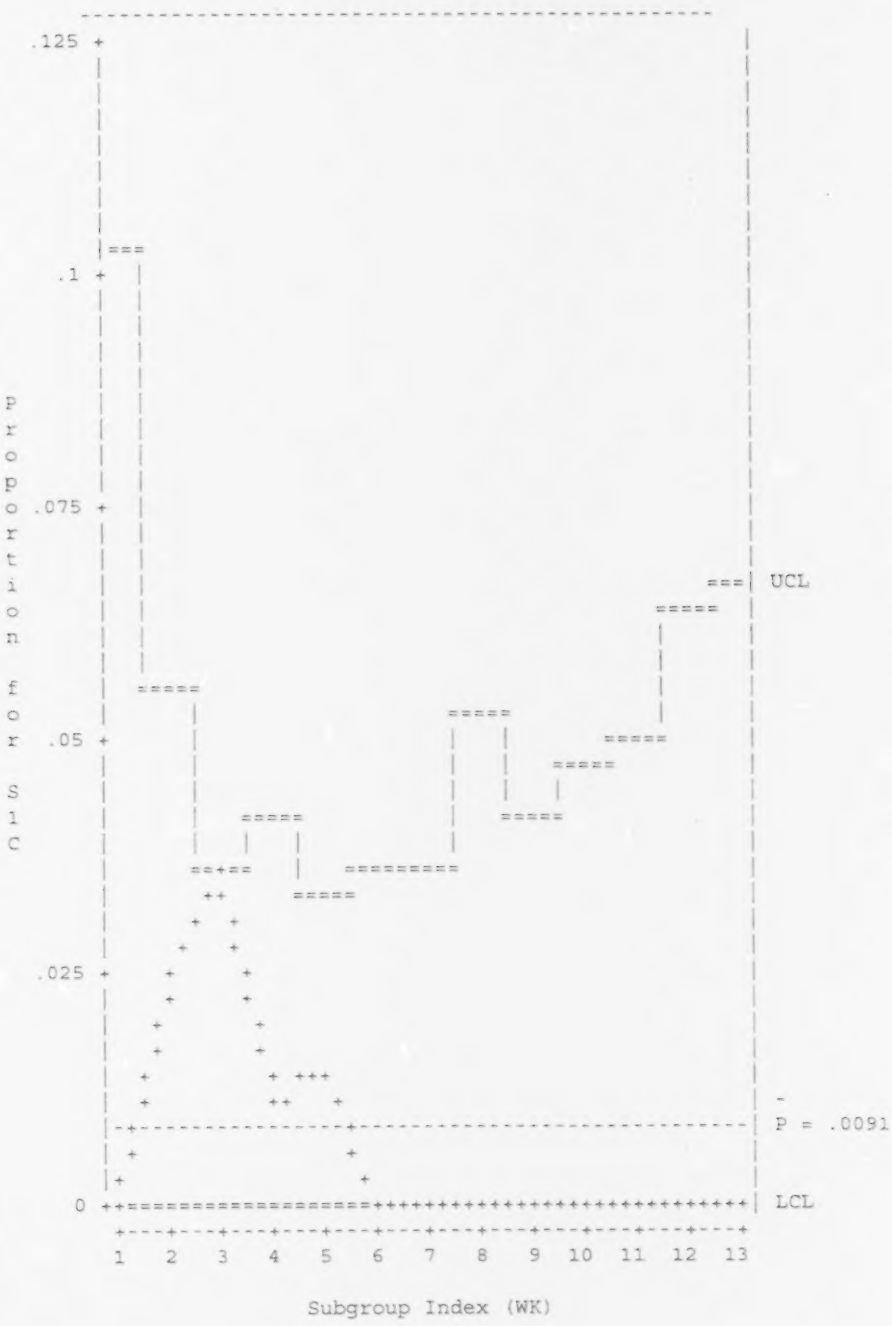
Figure 2.--Problems with item S1B, reads all appropriate phrases and answer categories



Subgroup Sizes: Min n=75 Max n=758

BEST COPY AVAILABLE

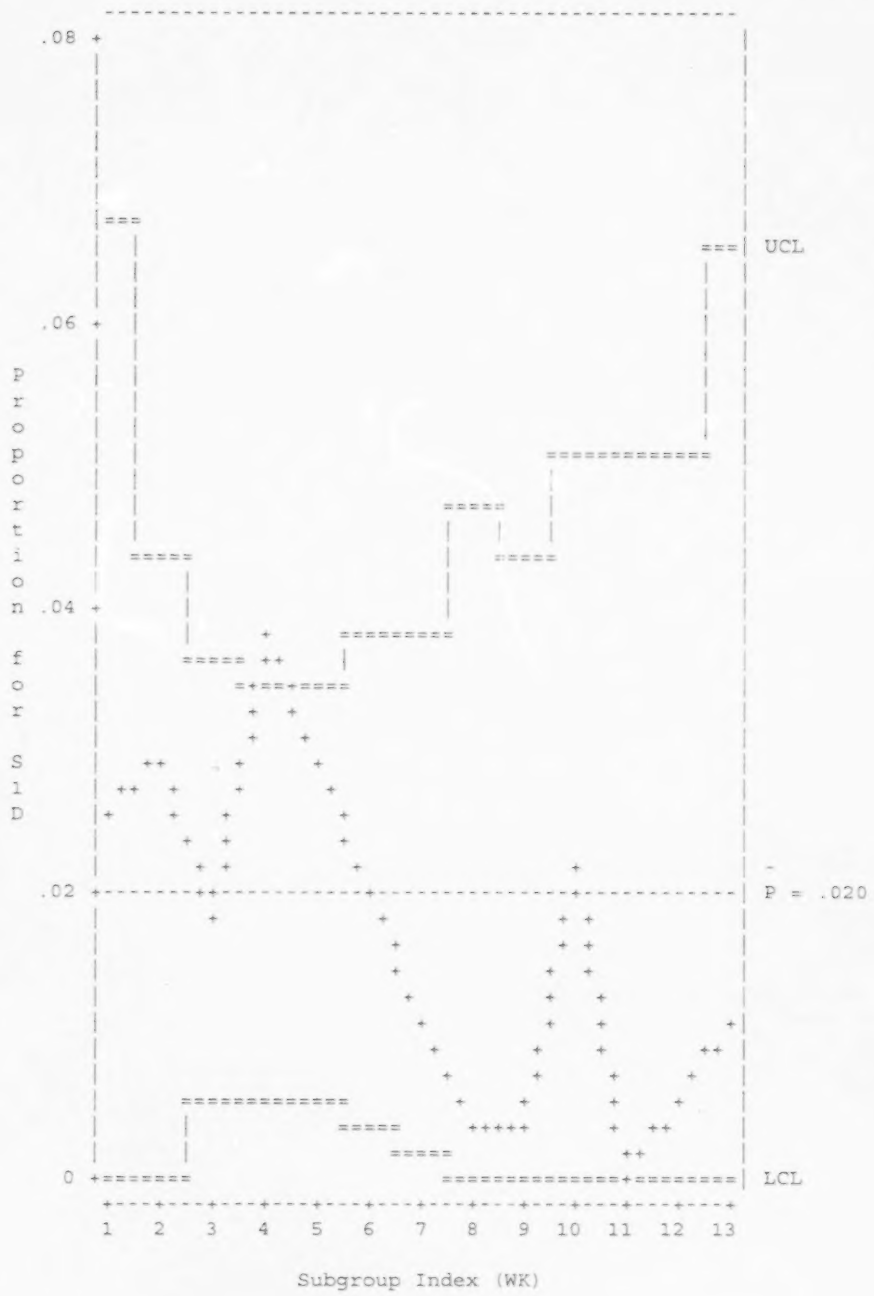
Figure 3.--Problems with item S1C, follows skip and box instructions



Subgroup Sizes: Min n=9 Max n=139

BEST COPY AVAILABLE

Figure 4.--Problems with item S1D, reads questions clearly with appropriate volume



Subgroup Sizes: Min n=78 Max n=802

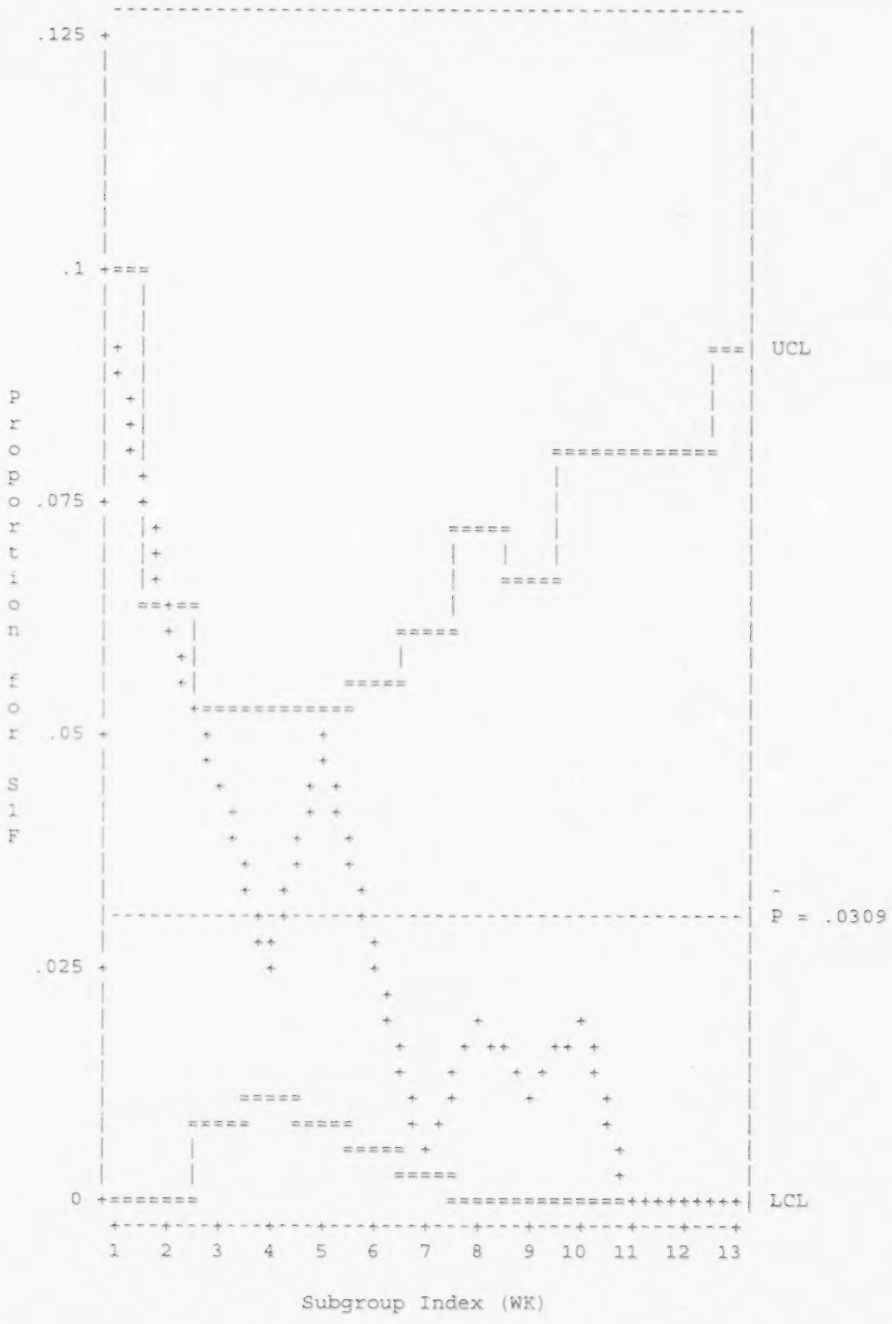
BEST COPY AVAILABLE

Figure 5.--Problems with item S1E, verifies spelling, address, phone numbers, and other data as needed

Subgroup Sizes: Min n=40 Max n=432

BEST COPY AVAILABLE

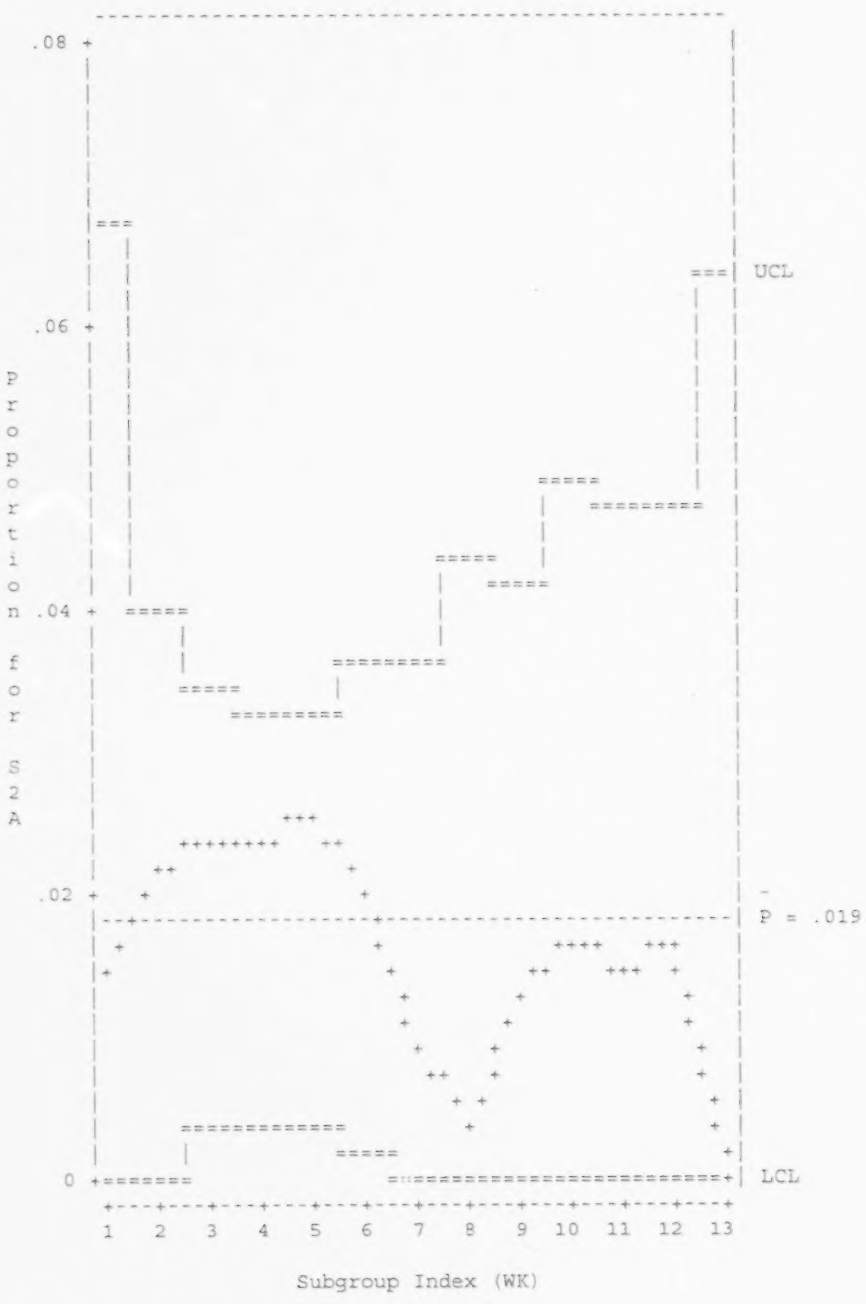
Figure 6.--Problems with item S1F, adjusts pace of interview to accomodate respondent



Subgroup Sizes: Min n=55 Max n=607

BEST COPY AVAILABLE

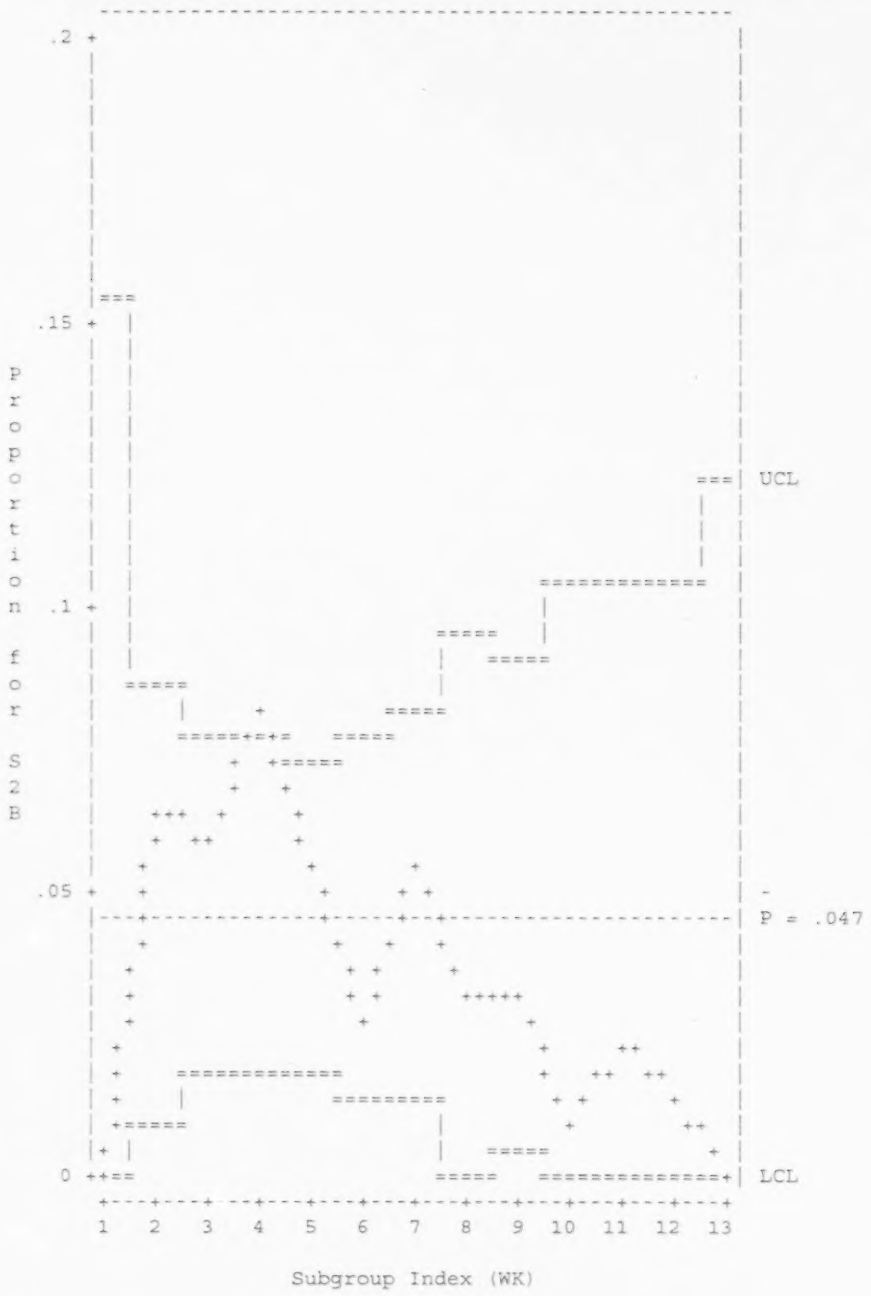
Figure 7.--Problems with item S2A, listens to entire answer



Subgroup Sizes: Min n=71 Max n=725

BEST COPY AVAILABLE

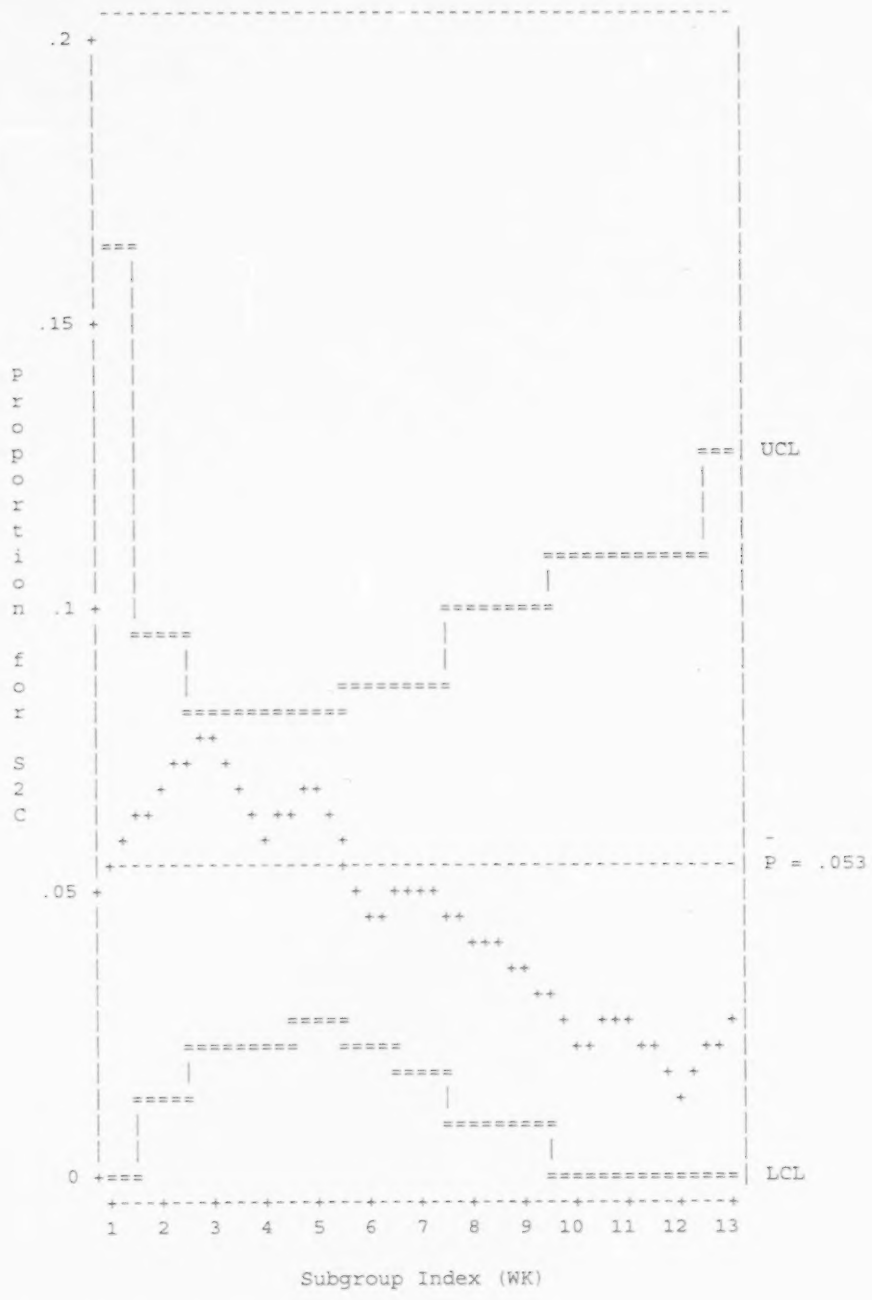
Figure 8.--Problems with item S2B, listens to what may not be said and probes



Subgroup Sizes: Min n=36 Max n=524

BEST COPY AVAILABLE

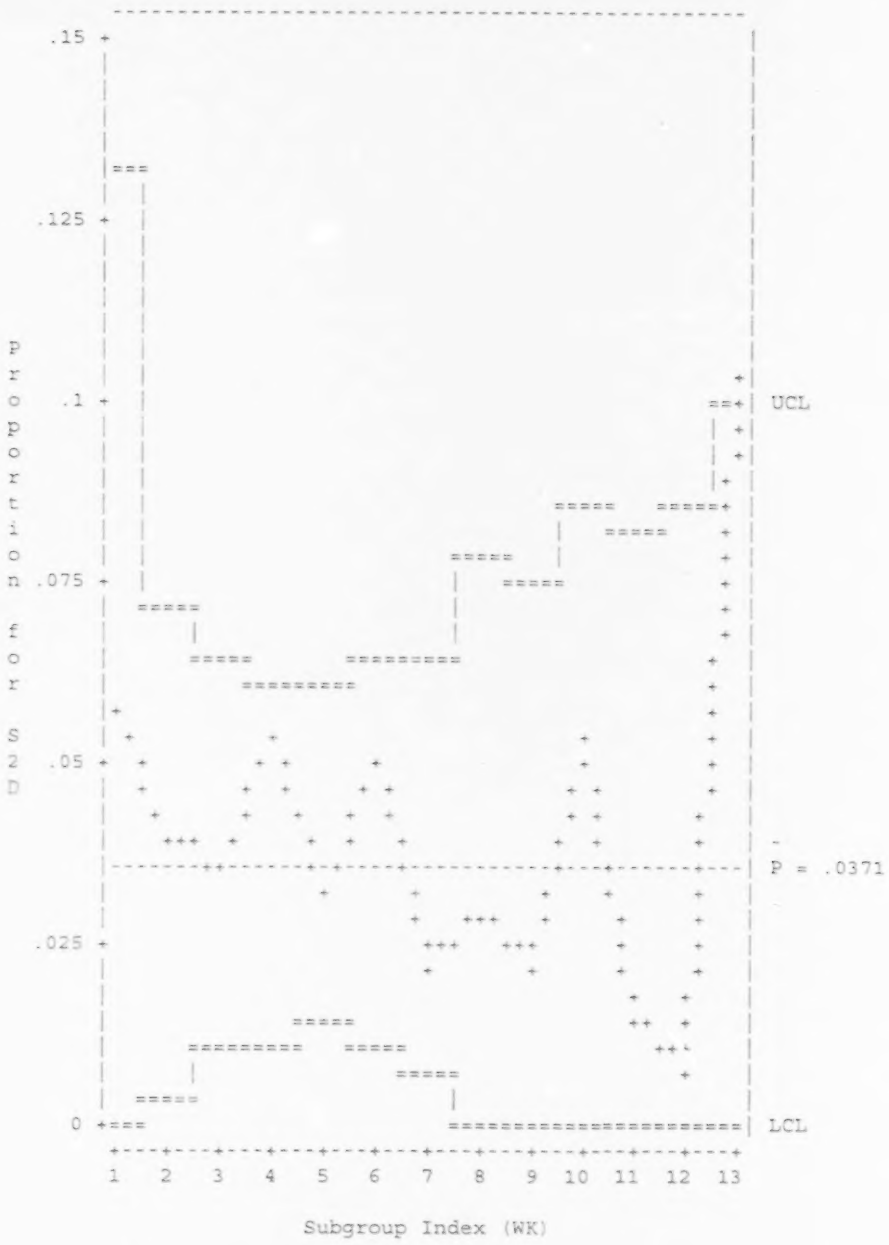
Figure 9.--Problems with item S2C, probes unclear responses



Subgroup Sizes: Min n=36 Max n=571

BEST COPY AVAILABLE

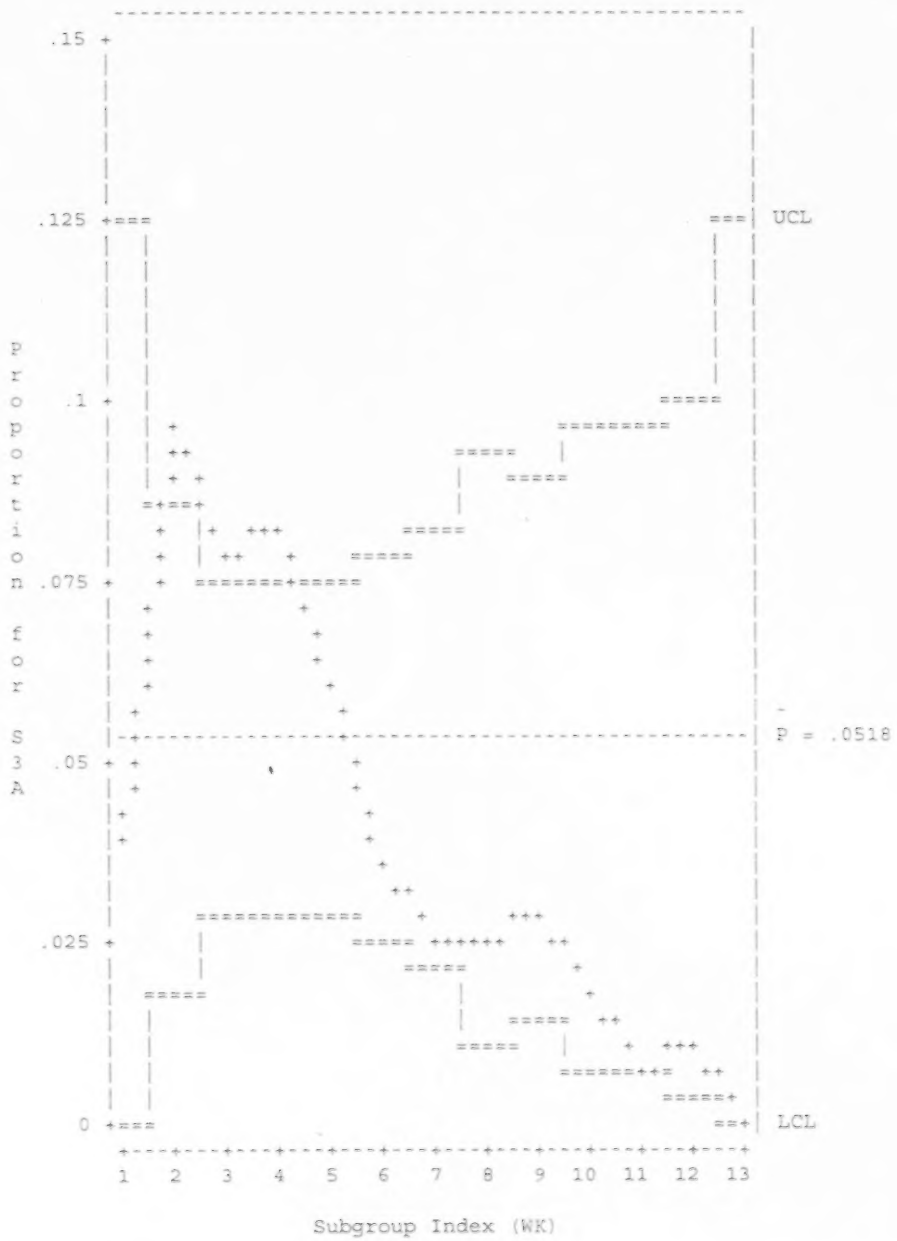
Figure 10.--Problems with item S2D, remains neutral when probing



Subgroup Sizes: Min n=36 Max n=557

BEST COPY AVAILABLE

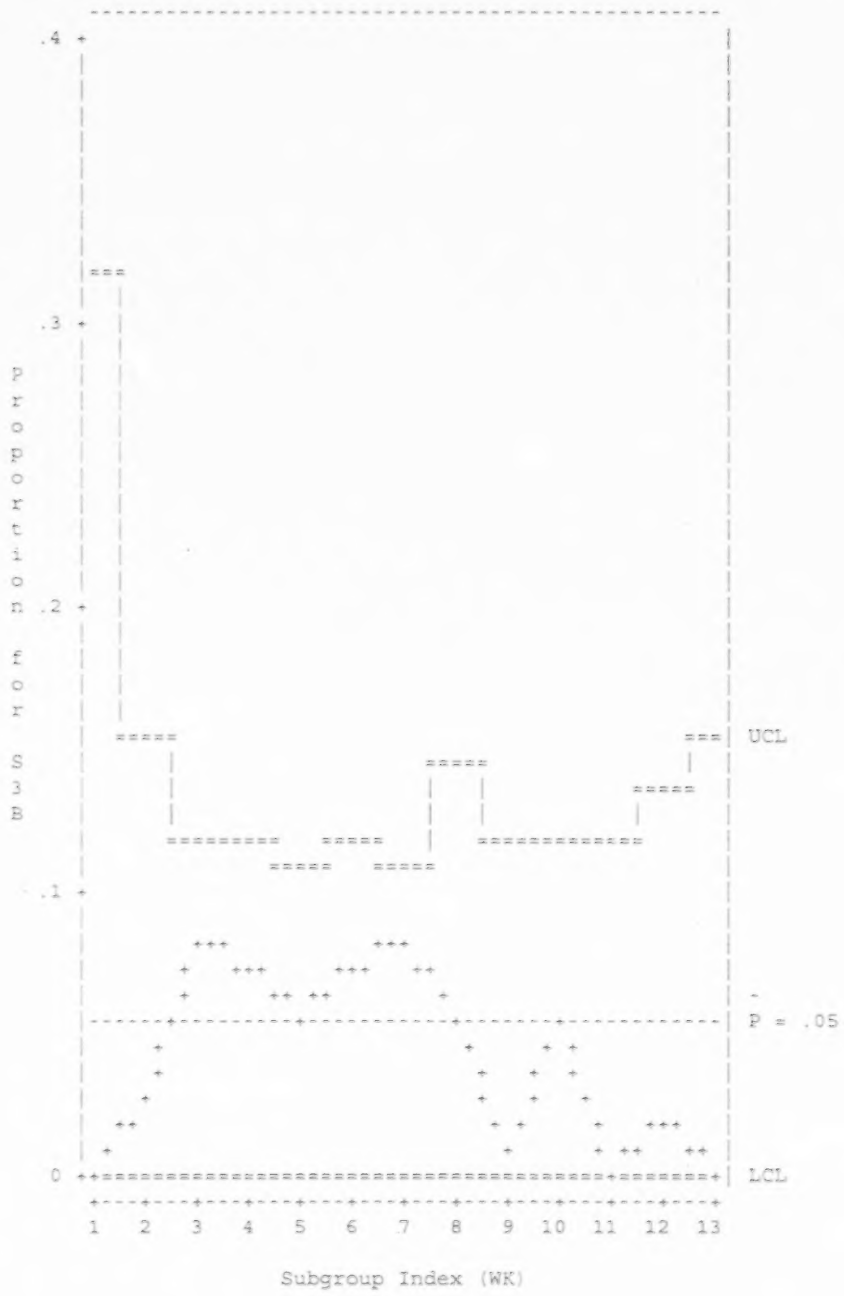
Figure 11.--Problems with item S3A, records information accurately



Subgroup Sizes: Min n=79 Max n=764

BEST COPY AVAILABLE

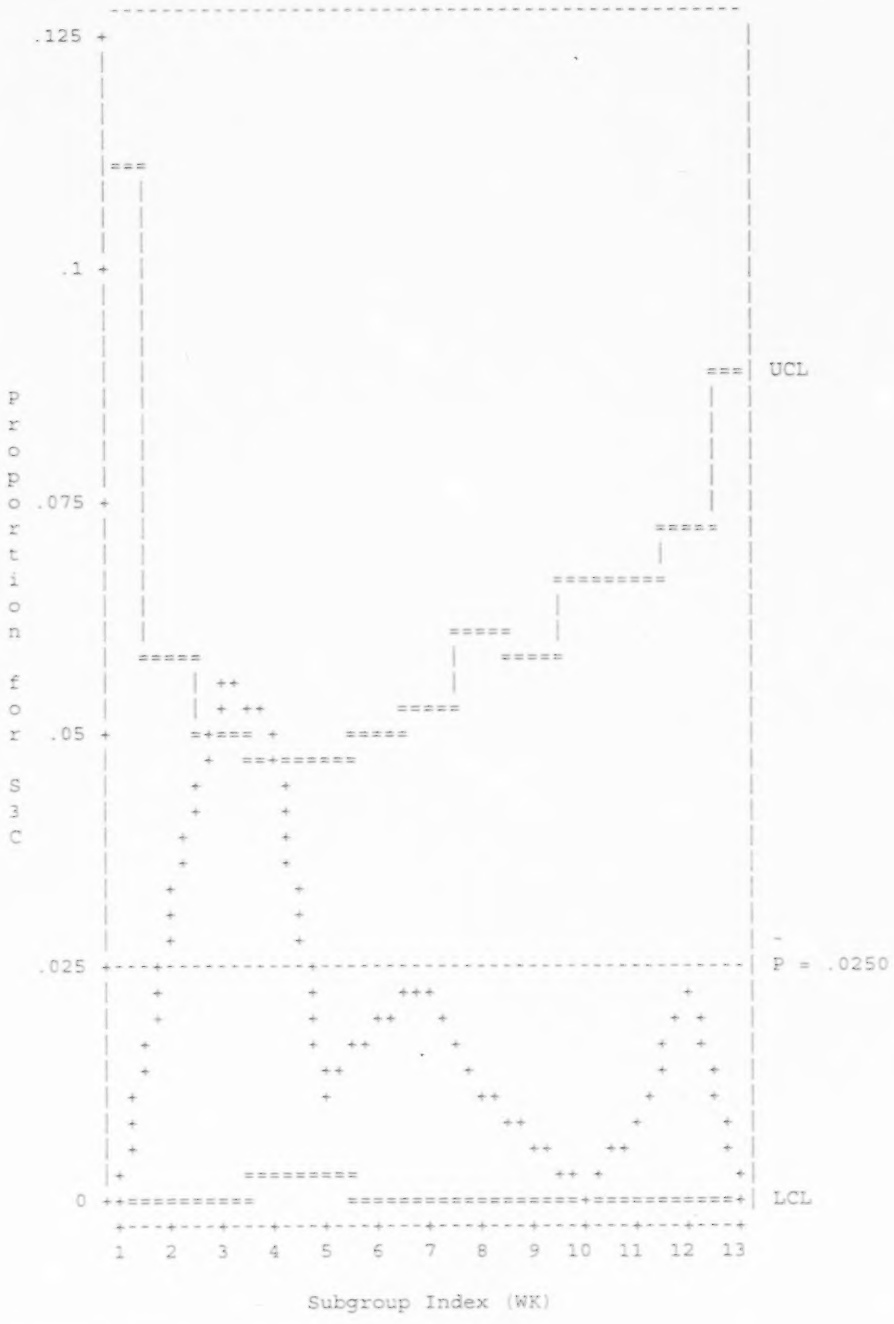
Figure 12.--Problems with item S3B, uses comments appropriately



Subgroup Sizes: Min n=6 Max n=137

BEST COPY AVAILABLE

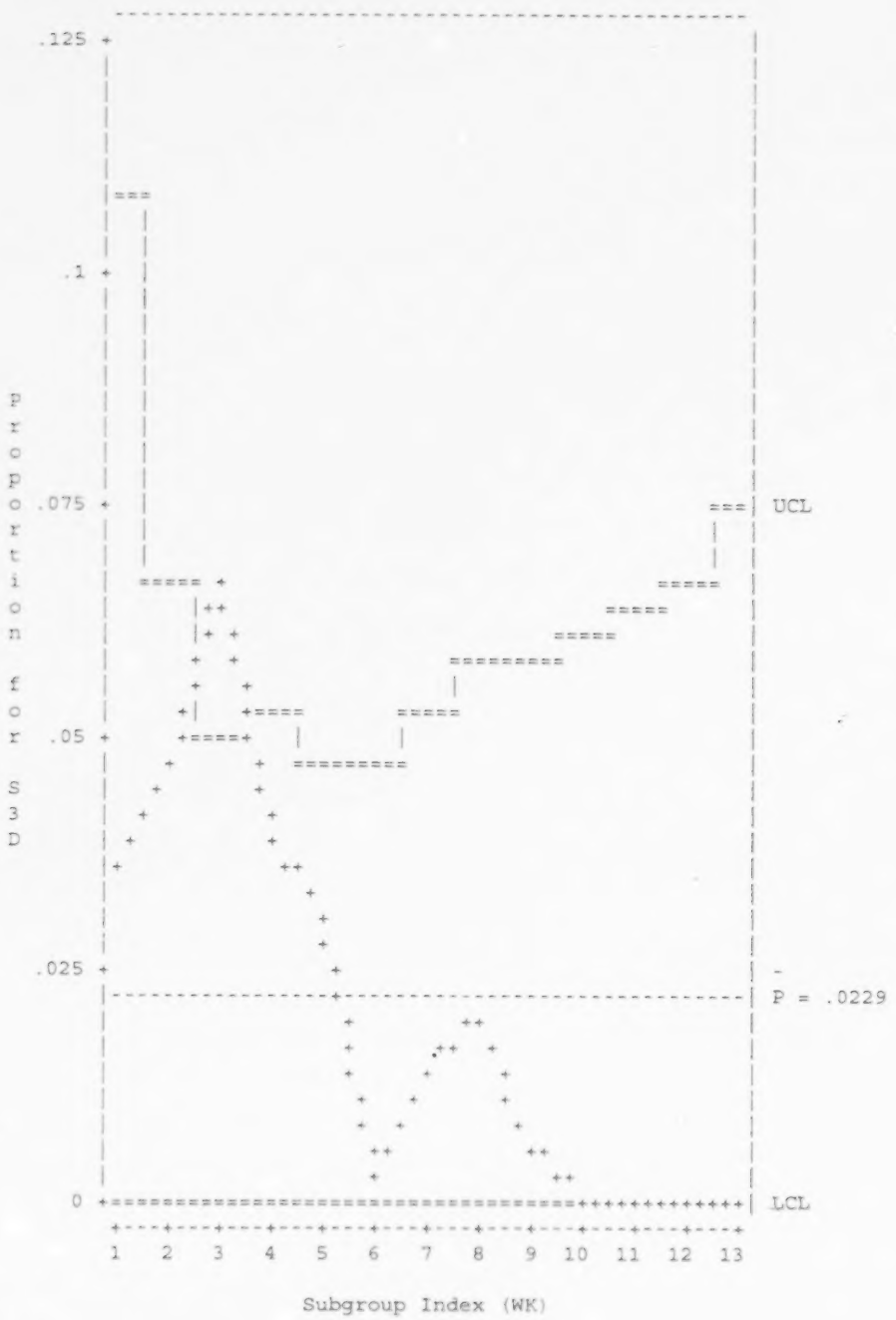
Figure 13.--Problems with item S3C, correct coding errors



Subgroup Sizes: Min n=30 Max n=447

BEST COPY AVAILABLE

Figure 14.--Problems with item S3D, uses control keys properly



Subgroup Sizes: Min n=27 Max n=320

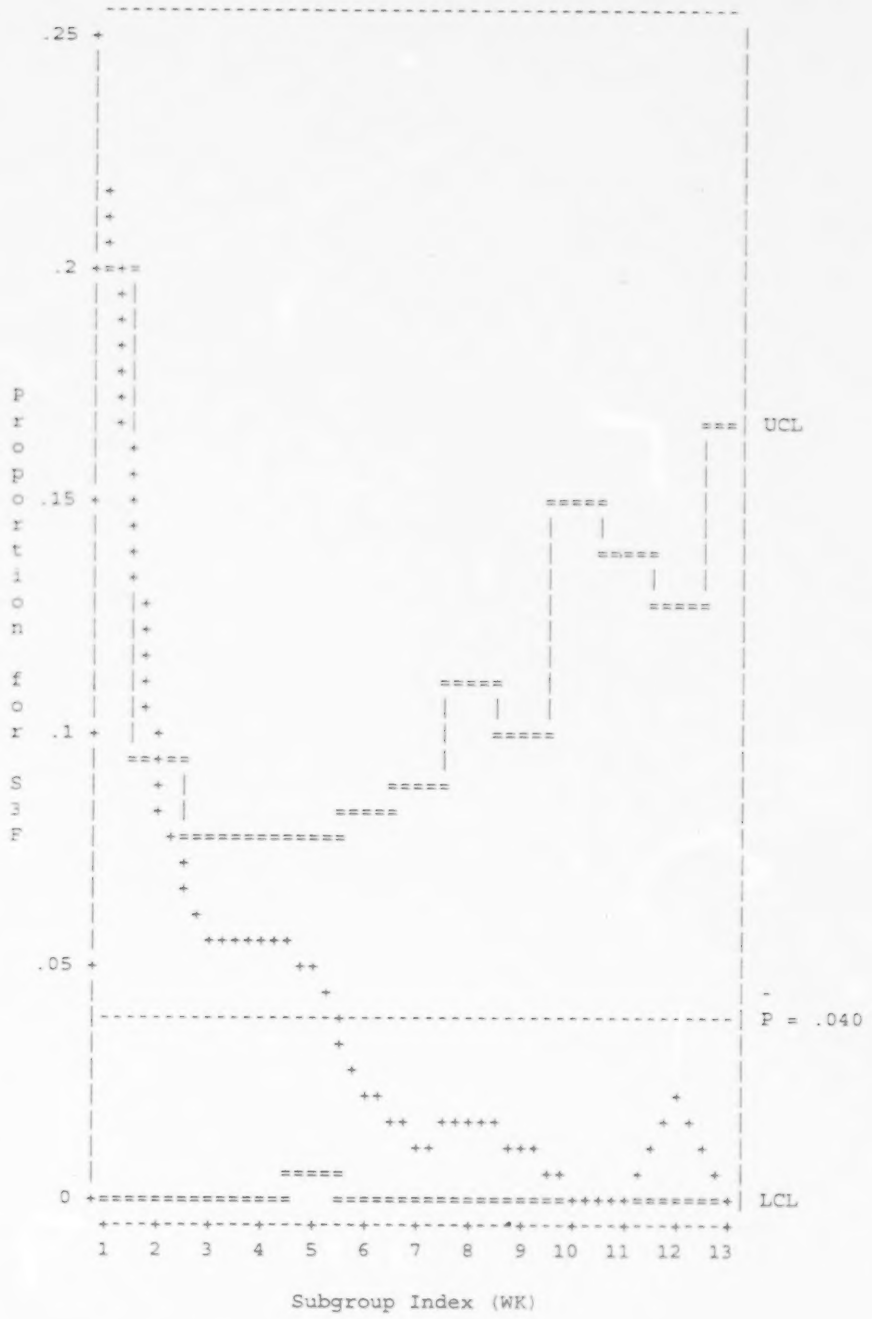
BEST COPY AVAILABLE

Figure 15.--Problems with item S3E, records result codes correctly

Subgroup Sizes: Min n=59 Max n=576

BEST COPY AVAILABLE

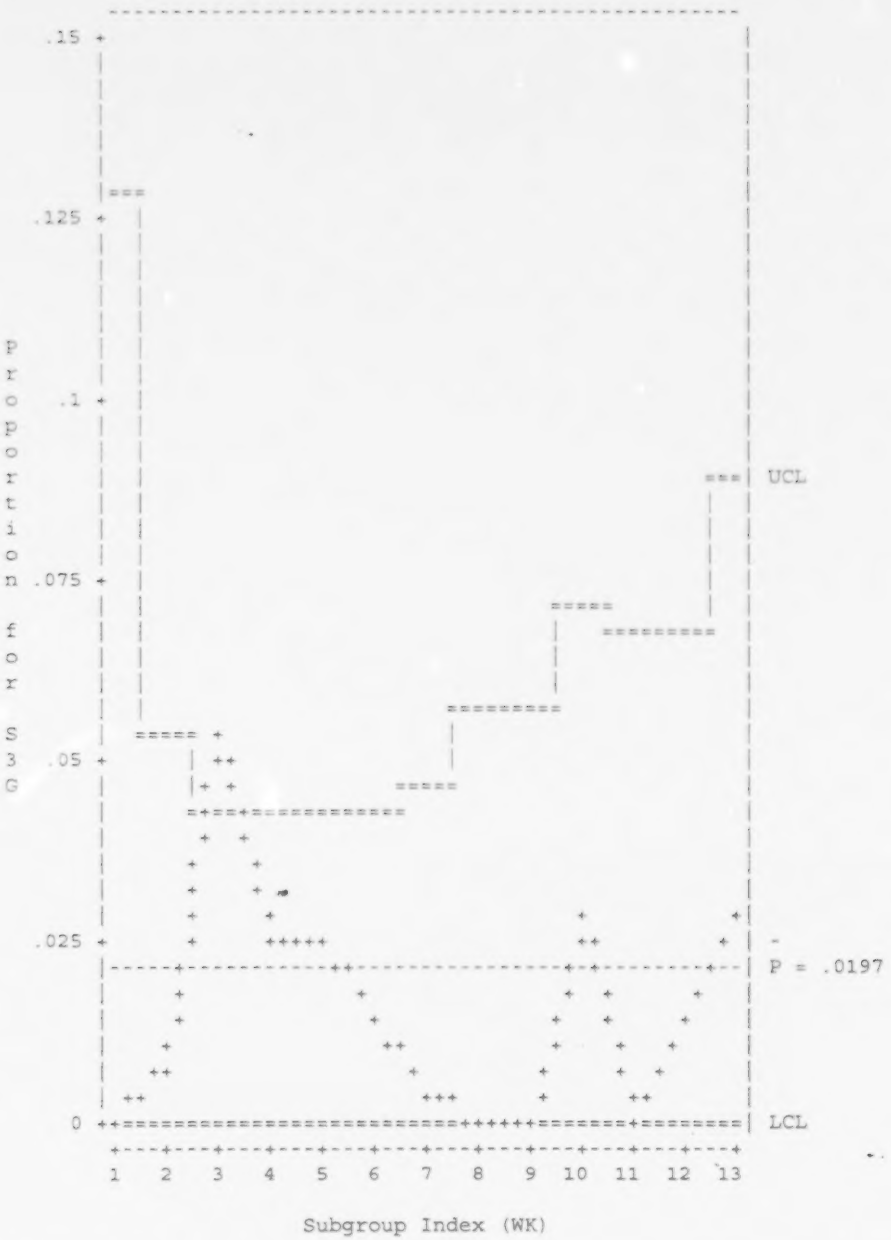
Figure 16.--Problems with item S3F, moved through matrix and selection screens properly



Subgroup Sizes: Min n=14 Max n=275

BEST COPY AVAILABLE

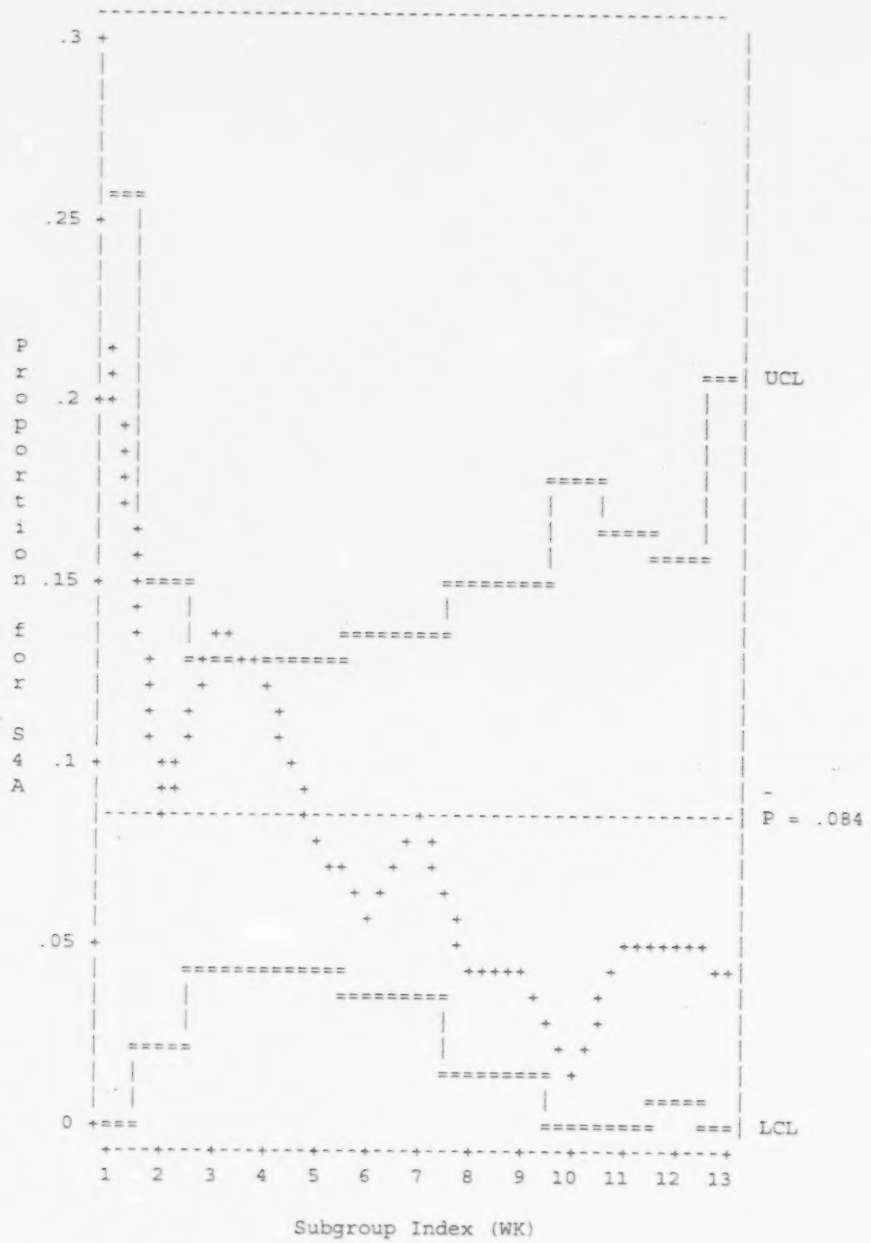
Figure 17.--Problems with item S3G, uses HHSELECT screens properly and smoothly



Subgroup Sizes: Min n=15 Max n=364

BEST COPY AVAILABLE

Figure 18.--Problems with item S4A, answers respondent questions and objections clearly, confidently and briefly without hesitation



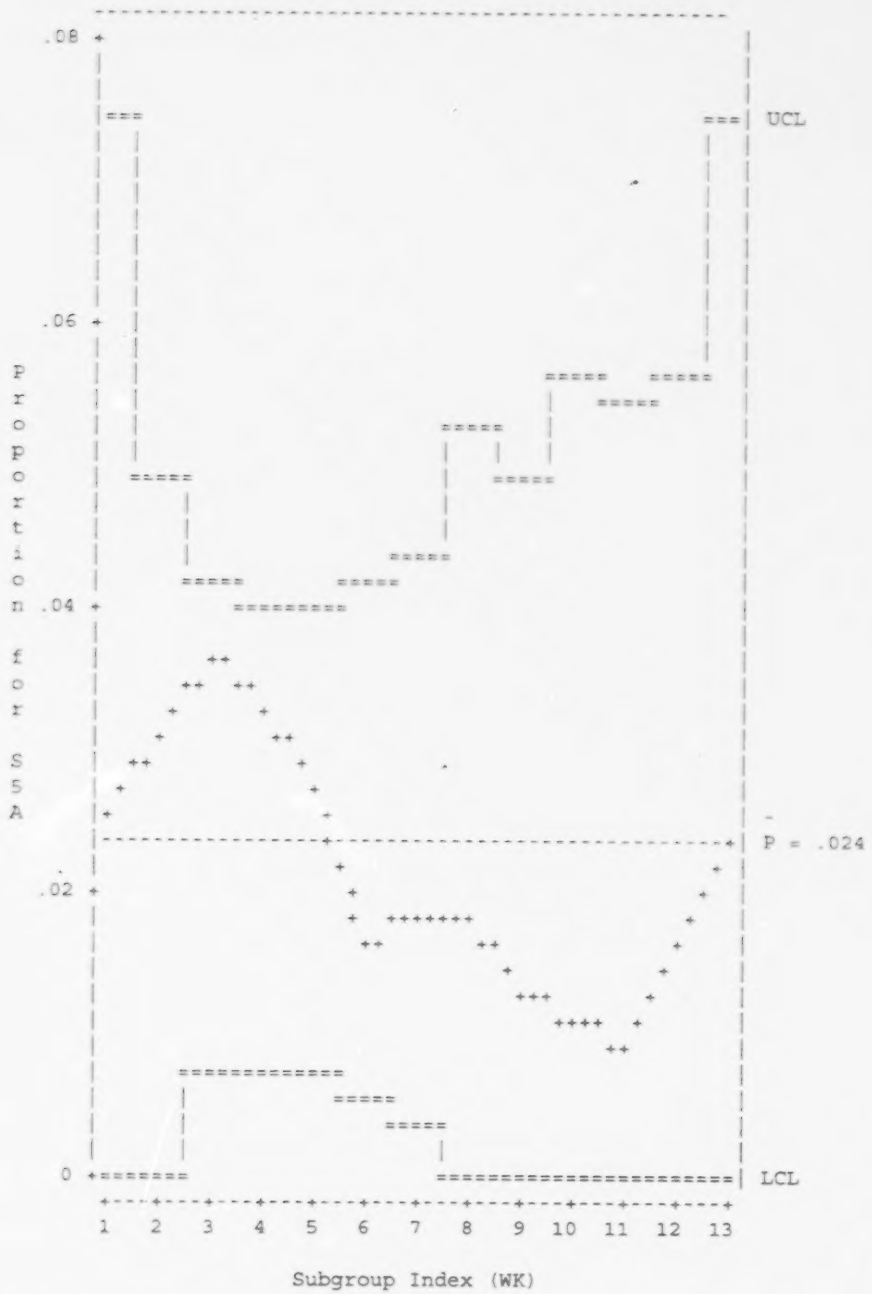
Subgroup Sizes: Min n=23 Max n=390

Figure 19.--Problems with item S4B, offers verification number

Subgroup Sizes: Min n=2 Max n=45

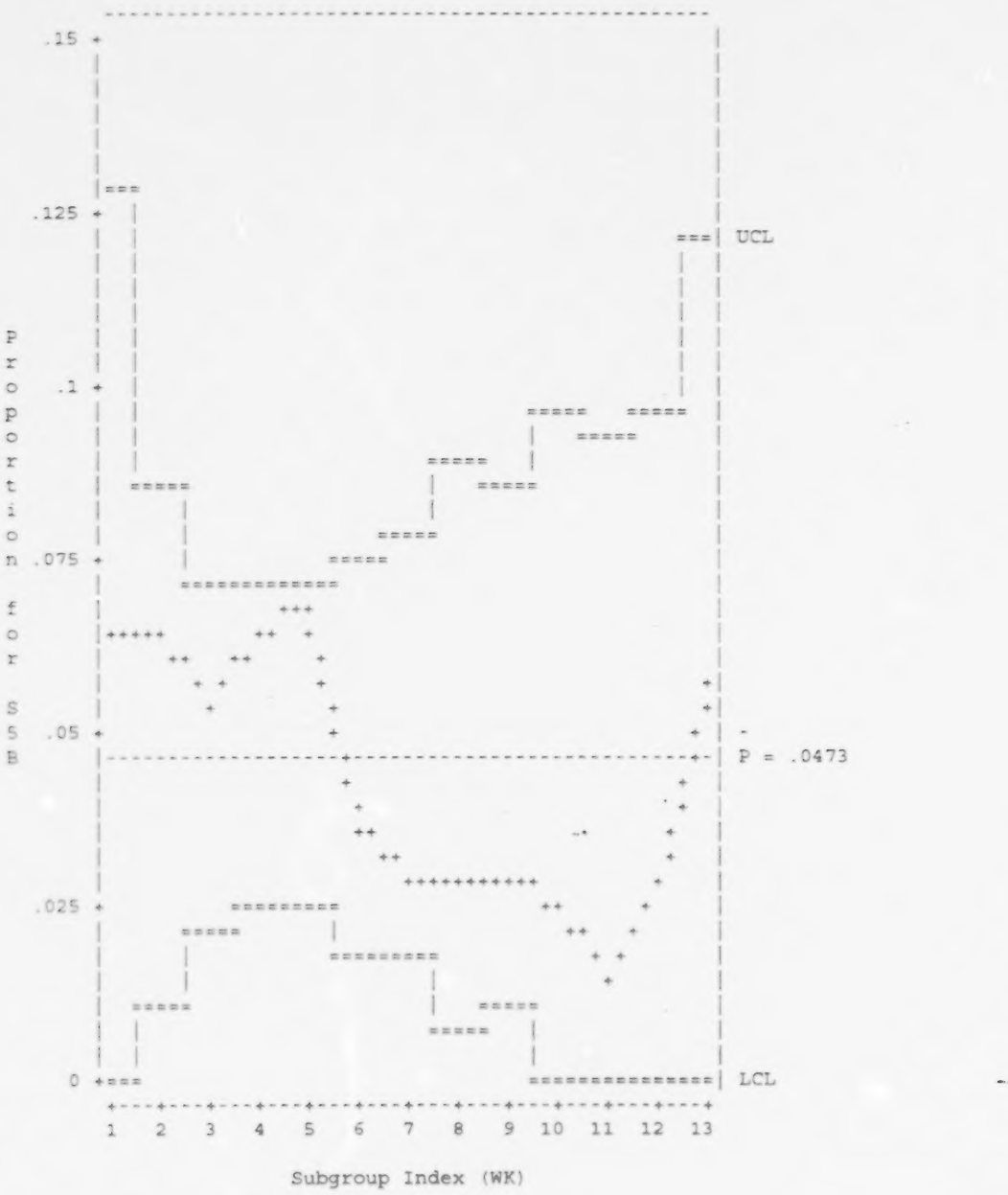
BEST COPY AVAILABLE

Figure 20.--Problems with item S5A, is pleasant, confident, and professional



Subgroup Sizes: Min n=80 Max n=812

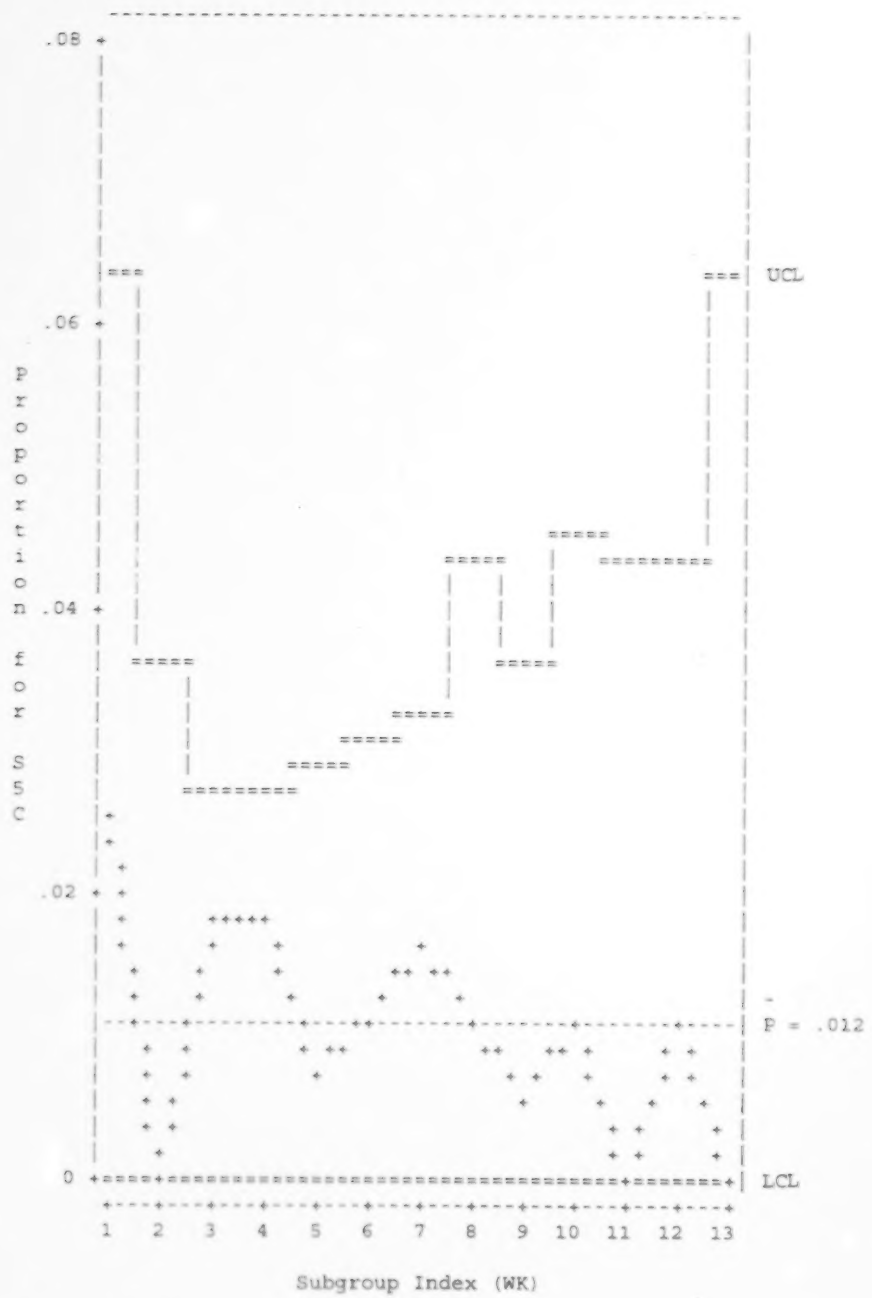
Figure 21.--Problems with item S5B, refrains from giving personal remarks or opinions



Subgroup Sizes: Min n=64 Max n=719

BEST COPY AVAILABLE

Figure 22.--Problems with item S5C, accepts emotions and sentiments without becoming personally involved



Subgroup Sizes: Min n=39 Max n=432

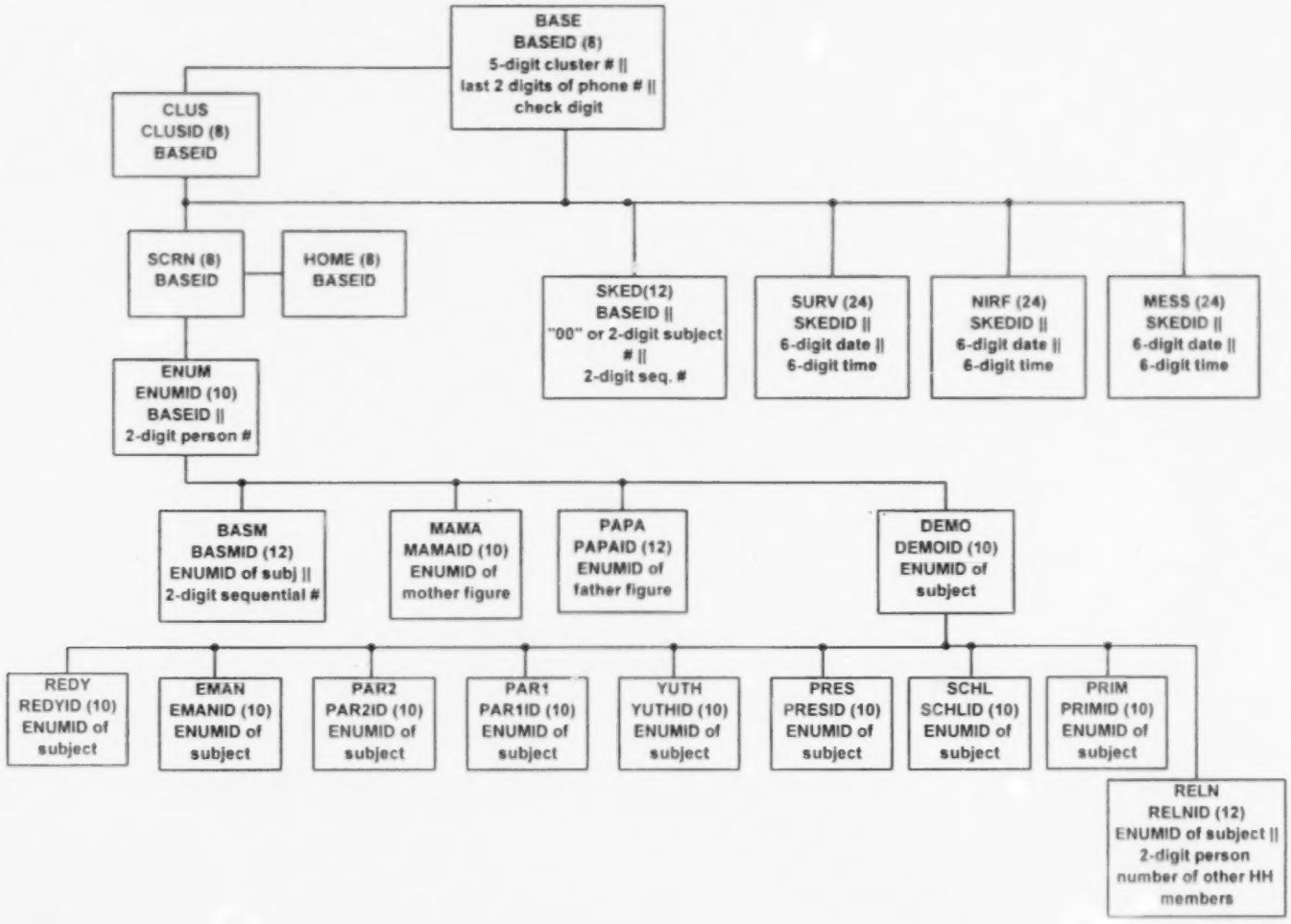
BEST COPY AVAILABLE

APPENDIX C

DATABASE DESIGN DIAGRAM

BEST COPY AVAILABLE

NHES: 93 Database Design Diagram



BEST COPY AVAILABLE

Listing of NCES Working Papers to Date

Please contact Ruth R. Harris at (202) 219-1831
if you are interested in any of the following papers

Number	Title	Contact
94-01 (July)	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02 (July)	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03 (July)	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04 (July)	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-05 (July)	Cost-of-Education Differentials Across the States	William Fowler
94-06 (July)	Six Papers on Teachers from the 1990-91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
94-07 (Nov.)	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
95-01 (Jan.)	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02 (Jan.)	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03 (Jan.)	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95-04 (Jan.)	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
95-06 (Jan.)	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-08 (Feb.)	CCD Adjustment to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09 (Feb.)	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10 (Feb.)	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11 (Mar.)	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12 (Mar.)	Rural Education Data User's Guide	Samuel Peng
95-13 (Mar.)	Assessing Students with Disabilities and Limited English Proficiency	James Houser
95-14 (Mar.)	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95-15 (Apr.)	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95-16 (Apr.)	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-17 (May)	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
95-18 (Nov.)	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96-01 (Jan.)	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-02 (Feb.)	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96-03 (Feb.)	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
96-04 (Feb.)	Census Mapping Project/School District Data Book	Tai Phan
96-05 (Feb.)	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96-06 (Mar.)	The Schools and Staffing Survey (SASS) for 1998-99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96-07 (Mar.)	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96-08 (Apr.)	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96-09 (Apr.)	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998-99 SASS	Dan Kasprzyk
96-10 (Apr.)	1998-99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96-11 (June)	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96-12 (June)	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96-13 (June)	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96-14 (June)	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-15 (June)	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk
96-16 (June)	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96-17 (July)	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
96-18 (Aug.)	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
96-19 (Oct.)	Assessment and Analysis of School-Level Expenditures	William Fowler
96-20 (Oct.)	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-21 (Oct.)	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96-22 (Oct.)	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96-23 (Oct.)	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96-24 (Oct.)	National Assessments of Teacher Quality	Dan Kasprzyk
96-25 (Oct.)	Measures of Inservice Professional Development: Suggested Items for the 1998-1999 Schools and Staffing Survey	Dan Kasprzyk
96-26 (Nov.)	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96-27 (Nov.)	Intersurvey Consistency in NCES Private School Surveys for 1993-94	Steven Kaufman

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-28 (Nov.)	Student Learning, Teaching Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection	Mary Rollefson
96-29 (Nov.)	Undercoverage Bias in Estimates of Characteristics of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
96-30 (Dec.)	Comparison of Estimates from the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97-01 (Feb.)	Selected Papers on Education Surveys: Papers Presented at the 1996 Meeting of the American Statistical Association	Dan Kasprzyk
97-02 (Feb.)	Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-03 (Feb.)	1991 and 1995 National Household Education Survey Questionnaires: NHES:91 Screener, NHES:91 Adult Education, NHES:95 Basic Screener, and NHES:95 Adult Education	Kathryn Chandler
97-04 (Feb.)	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler

BEST COPY AVAILABLE